ABSTRACT

Title of dissertation: SPATIAL AND TEMPORAL MODELING OF LARGE-SCALE BRAIN NETWORKS

Mahshid Najafi,
Doctor of Philosophy, 2017

Dissertation directed by: Professor Jonathan Z. Simon,
Department of Electrical and Computer Engineering
Professor Luiz Pessoa,
Department of Psychology

The human brain is the most fascinating and complex organ. It directs all our actions and thoughts. Despite the large body of brain studies, little is known about the neural basis of its large-scale structure. In this dissertation, I take advantage of several network-based and statistical techniques to investigate the spatial and temporal aspects of large-scale functional networks of the human brain during "rest" and "task" conditions using functional MRI data.

Large-scale analysis of human brain function has revealed that brain regions can be grouped into networks or communities. Most studies adopt a framework in which brain regions belong to only one community. Yet studies in general fields of knowledge suggest that in most cases complex networks consist of interwoven sets of overlapping communities. A mixed-membership framework can better characterize the complex networks. In this dissertation, I employed a mixed-membership Bayesian model to characterize overlapping community structure of the brain at

both “rest” and “task” conditions. The approach allowed us to quantify how task performance reconfigures brain communities at rest, and determine the relationship between functional diversity (how diverse is a region’s functional activation repertoire) and membership diversity (how diverse is a region’s affiliation to communities). Furthermore, I could study the distribution of key regions, named “bridges”, in transferring information across the brain communities. Our findings revealed that the overlapping framework described the brain in ways that were not captured by disjoint clustering, and thus provided a richer landscape of large-scale brain networks. Overall, I suggest that overlapping networks are better suited to capture the flexible and task-dependent mapping between brain regions and their functions.

Finally, I developed a dynamic intersubject network analysis technique to study the temporal changes of the emotional brain at the level of large-scale brain networks by formulating a manipulation in which threat levels varied continuously during the experiment. Our results illustrate that cohesion within and between networks changed dynamically with threat level. Together, our findings reveal that characterizing emotional processing should be done at the level of distributed networks, and not simply at the level of evoked responses in specific brain regions.

SPATIAL AND TEMPORAL MODELING OF LARGE-SCALE BRAIN NETWORKS

by

Mahshid Najafi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2017

Advisory Committee:
Professor Jonathan Z. Simon, Chair/Advisor
Professor Luiz Pessoa, Advisor
Professor Joseph JaJa
Professor Behtash Babadi
Professor Donald J. Bolger, Dean's Representative

Dedication

To my wonderful family, and in memory of Forough..

Acknowledgments

I would like to thank my advisor and mentor, Professor Luiz Pessoa, for his excellent guidance, support, feedback, and for always encouraging me to push my boundaries. I could never imagine having a better advisor than Luiz, and I will always be grateful for that. This dissertation wouldn't be done without all our long discussions in the office and over Skype which gave me insight on how to do the data analysis.

I also would like to thank my ECE advisor and mentor, Professor Jonathan Z. Simon, for all his great support and guidance. I'll never forget how helpful and supportive he was in difficult times.

I would like to thank Profs. Joseph JaJa, Behtash Babadi, and Donald J. Bolger for serving on the thesis committee and Dr. Piya Pal for serving on the research proposal committee.

I would like to acknowledge the Behavioral and Social Sciences College, University of Maryland for the high-performance computing resources (http://bsos.umd.edu/oacs/bsos-high-performance) made available for conducting the research reported in this dissertation.

I'm also deeply grateful to all the Pessoa Lab members (current and former) for their support, help, and for making the office environment a pleasant workplace as well as for being great friends and colleagues. I'd like to thank Srikanth Padmala, who helped me to become familiar with fMRI data analysis, and for always being there to answer all my questions; Christian for his help in English editing of this dissertation, data acquisition and preprocessing; Josh Kinnison for providing us

with his great code and toolboxes, and more importantly encouraging us to code in Python, and bringing his cute dog, Eco, around; Brenton for his bright comments and ideas; and Sandra and Moon for their support in my early days in the lab; Mihai, Dan, Jessica, Jennifer for being great research assistants; and Alessia for giving feedback on parts of the research proposal report.

No life can be meaningful without the laughter we share with our friends. I would like to thank all my friends who made wonderful moments in my life: Atefeh, Fatemeh Vali, Elham, Hadi, Bahar, Ladan, Niloofar, Alborz, Alireza, Pouya, Arian, Elnaz, Hamidreza, Faryaneh, Moein, Marjan, Mostafa, Sahar, Mahyar, Shohreh, Helia, Shahrzad, Ghazaal, Amirhosseion, Ehsan, Miorel, Victor, Julius, Jennifer Stark, Jason Smith, and all my other dear friends and my cousins.

I would like to thank my parents and my brother for their lifetime support, endless kindness, and for always being there for me and encouraging me to follow my dreams. I wish they could come here to see this happening.

I would like to say my thanks to my wonderful husband, love, and best friend, Ali. I'm indebted to him for his endless support, being there for me in every moment of my graduate school, and making our house a place of happiness and peace. I'm so lucky to have him in my life.

Mahshid Najafi

Spring 2017

Contents

List of Tables viii

List of Figures ix

1 Introduction 1
1.1 The Brain: A large-scale complex system . . . 1
1.2 Thesis outline . . . 2

2 Background 7
2.1 Brain Networks . . . 7
2.1.1 Structural network . . . 7
2.1.2 Functional network . . . 9
2.1.3 Effective network . . . 11
2.1.4 Brain Communities . . . 12
2.2 functional Magnetic Resonance Imaging . . . 15
2.2.1 Principles of fMRI . . . 15
2.2.2 MRI scanner and BOLD Imaging . . . 16
2.2.3 MRI data preprocessing pipeline . . . 17

3 Overlapping communities reveal rich structure in large-scale brain networks during rest 20
3.1 Introduction . . . 20
3.2 Materials and Methods . . . 24
3.2.1 Resting-state data . . . 24
3.2.2 Regions of Interest . . . 25
3.2.3 Functional connectivity . . . 26
3.2.4 Detecting disjoint communities . . . 27
3.2.5 Detecting overlapping communities . . . 27
3.2.5.1 Mixed-membership algorithm . . . 28
3.2.5.2 Choosing the number of communities . . . 31
3.2.6 Reproducibility and reliability of the results . . . 31
3.2.6.1 Community similarity . . . 33
3.2.7 Functional diversity . . . 33
3.2.8 Membership diversity . . . 35
3.2.9 Cortical and Subcortical Visualization . . . 35
3.3 Results . . . 36

3.3.1 Comparing disjoint and overlapping communities at rest . . . 37
3.3.2 Relationship between functional diversity and membership diversity . . . 40
3.3.3 Reliability of results . . . 42
3.3.4 Does overlapping community structure at the group level reflect that at the participant level? . . . 43
3.4 Discussion . . . 48

4 Overlapping communities revealed important differences between community organization during rest and performing task states 53
4.1 Introduction . . . 53
4.2 Materials and Methods . . . 55
4.2.1 Datasets . . . 55
4.2.1.1 Working memory task data . . . 55
4.2.1.2 Emotion task data . . . 56
4.2.1.3 Data censoring . . . 57
4.2.2 Regions of Interest . . . 58
4.2.3 Functional connectivity . . . 58
4.2.4 Detecting disjoint communities . . . 58
4.2.5 Detecting overlapping communities . . . 59
4.2.6 Degree . . . 61
4.2.7 Modularity of overlapping networks . . . 62
4.2.8 Defining Bridges . . . 63
4.3 Results . . . 64
4.3.1 Comparing overlapping communities during rest and task . . . 64
4.3.2 Disjoint communities during the working memory task . . . 66
4.3.3 Modularity of overlapping networks . . . 67
4.3.4 Node taxonomy: hubs and bridges during rest and task conditions . . . 69
4.3.5 Reliability of results . . . 73
4.4 Discussion . . . 75
4.4.1 How do tasks alter the functional connectivity landscape? . . . 76
4.4.2 Characterizing different types of bridges . . . 78

5 Intersubject brain network organization during dynamic anxious anticipation 82
5.1 Introduction . . . 82
5.2 Materials and Methods . . . 86
5.2.1 Data and Experiment Design . . . 86
5.2.2 Functional MRI preprocessing . . . 90
5.2.3 Exploratory and test sets . . . 93
5.2.4 Regions of interest . . . 94
5.2.5 Intersubject functional network . . . 97
5.2.6 Dynamic intersubject networks . . . 99
5.2.7 Within- and between-network cohesion . . . 100
5.3 Results . . . 102

5.3.1 Intersubject network analysis and statistical approach 102
5.3.2 Network cohesion 104
5.3.3 Subcortical regions. 106
5.3.4 Dynamics 108
5.4 Discussion 108

6 Conclusions and future directions 115
6.1 The Brain: An interwoven network 115
6.2 The emotional brain: from a dynamic large-scale network perspective 119
6.3 Future directions 121

A Supplemental Materials of Chapter 3 123
A.1 supplemental materials 123
A.1.1 Membership diversity and participation coefficient 123
A.2 Supplemental Figures 124

B Supplemental Materials of Chapter 4 130
B.1 supplemental materials 130
B.1.1 Potential impact of preprocessing pipelines on resting-state and task results 130
B.1.2 Potential impact of community size on bridge type 132
B.1.3 Potential impact of number of communities on modularity scores 134
B.2 Supplemental Figures 135
Bibliography 139

List of Tables

3.1 Correlation between membership diversity at the group and subject level. Only regions with nonzero membership value for communities OC1-OC6 were considered. 46

5.1 **List of Cortical Regions of interest.** Cortical ROIs were defined via 5-mm radius spheres centered on MNI coordinates provided below. 96

5.2 **List of Subcortical Regions of Interest.** Subcortical ROIs were defined anatomically (see text for details). Numbering is continued from Table 5.1 97

List of Figures

2.1 The 7-network cortical parcellation of resting-state fMRI data by Yeo et al. (2011). Based on anatomical and functional considerations, the communities were labeled as "Visual", "Somatomotor", "Ventral attention", "Dorsal attention" , "Limbic", "Frontoparietal", and "Default". The figure is adopted from Yeo et al. (2011). 14

3.1 Community organization in the brain. (A) Standard communities are disjoint (inset: colors indicate separate communities), as illustrated via three types of representation. The representation on the right corresponds to a schematic correlation matrix. (B) Overlapping communities are interdigitated, such that brain regions belong to multiple communities simultaneously (inset: community overlap shown on the brain indicated by intersection colors). 22

3.2 Overlapping communities and membership values. Each brain region affiliates with each community with varying strengths that are captured by the membership value. These probability-like values are between 0 and 1 and sum to 1 (for each region of interest, or ROI). . 23

3.3 Graphical representation of the mixed-membership model. Green circles indicate observed variables; orange circles indicate hidden variables. Box indicates replication over communities. Based on Gopalan and Blei (2013). 29

3.4 Model fit as a function of the number of communities for resting-state dataset investigated. Because the trend continues after $k = 20$, we do not show values $20 < k \leq 25$. 32

3.5 Disjoint communities (DC) detected during rest. Each row depicts one of the six disjoint communities extracted with k-means. Cortical and subcortical regions belonging to each of community are colored in red. 38

3.6 Overlapping communities (OC) detected during rest. Each row depicts one of the six overlapping communities extracted with the mixed-membership model. The color of cortical and subcortical regions reflects the median membership value of each region to each community across 5,000 iterations. Membership values are thresholded at 0.1 for illustration. 39

3.7 Cosine similarity between disjoint (DC1-DC6) and overlapping (OC1-OC6) communities during rest. The matrix displays the median cosine similarity between community pairs across 5,000 iterations. Given that community membership vectors do not contain negative values, the cosine similarity scores range from zero (orthogonal) to one (identical). Side and top insets represent similarity scores as bar plots across rows and columns, respectively. 40
3.8 Frequency histograms of membership values for each of the six overlapping communities during rest. Each histogram depicts the median value in each bin for that community across 5,000 iterations (error bars show the of 25th-75th percentile range). The colors of the bars correspond to the range of membership values shown in the brain insets. 41
3.9 Functional diversity scores (Shannon entropy) based on studies in the BrainMap database. Cool colors represent regions of low diversity and warm colors represent regions of high diversity. 42
3.10 Membership values for all ROIs with nonzero values reordered in ascending fashion during rest. Green: median; blue: 2.5th percentile; red: 97.5th percentile. 43
3.11 Overlapping community organization during rest at the individual level for sample participants. Each row depicts the six overlapping communities extracted with the mixed-membership model per subject (sOC). Membership values are thresholded at 0.1 for illustration. . . . 45
3.12 Frequency histograms of membership values for overlapping communities estimated at the individual level during rest. Bins show median values across 94 subjects (error bars show the of 25th-75th percentile range). Brain insets show median membership values across subjects (thresholded at 0.1 for illustration). The colors of the bars correspond to the range of membership values shown in the brain insets. 46
3.13 Membership diversity values are ranked at the group level and at the subject level, and each region's value is assigned its rank instead of its membership diversity value. The same is done for the average of subject level 47
3.14 Group-level membership diversity of all nodes vs. average of membership diversity at subject level (see also Figure S17). Each different colored node shows a region with robustly nonzero membership value in that community. 47

4.1 Model fit as a function of the number of communities for all conditions (resting-state dataset, working memory, and emotion) investigated. Because the trend continues after $k = 20$, we do not show values $20 < k \leq 25$. 60

4.2 Overlapping communities (WM) detected during the working memory task. Each row depicts one of the six overlapping communities extracted with the mixed-membership model. The color of cortical and subcortical regions reflects the median membership value of each region to each community across 5,000 iterations. Membership values are thresholded at 0.1 for illustration. 66
4.3 Overlapping Communities (EM) detected during the emotion task. Each row depicts one of the six overlapping communities extracted with the mixed-membership model. The color of cortical and subcortical regions reflects the median membership value of each region to each community across 5,000 iterations. Membership values are thresholded at 0.1 for illustration. 67
4.4 Cosine similarity (equation 3.1) between overlapping communities at rest (OC1-OC6) and during the working memory task (WM1-WM6). The matrix displays the median cosine similarity between each WM and OC network across 5,000 iterations. Given that the community membership vectors do not contain negative values, the cosine similarity scores range from zero (orthogonal) to one (identical). Side and top insets represent similarity scores as bar plots across rows and columns, respectively. 68
4.5 Cosine similarity (equation 3.1) between overlapping communities at rest (OC1-OC6) and during the emotion task (EM1-EM6). The matrix displays the median cosine similarity between each EM and OC network across 5,000 iterations. Given that the community membership vectors do not contain negative values, the cosine similarity scores range from zero (orthogonal) to one (identical). Side and top insets represent similarity scores as bar plots across rows and columns, respectively. 69
4.6 Frequency histograms of membership values for each of the six overlapping communities during the working memory task. Each histogram depicts the median value in each bin for that community across 5,000 iterations (error bars show the of 25th-75th percentile range). The colors of the bars correspond to the range of membership values shown in the brain insets. 70
4.7 Frequency histograms of membership values for each of the six overlapping communities during the emotion task. Each histogram depicts the median value in each bin for that community across 5,000 iterations (error bars show the of 25th-75th percentile range). The colors of the bars correspond to the range of membership values shown in the brain insets. 71
4.8 Disjoint communities during the working memory task estimated by the Infomap algorithm. 71
4.9 Disjoint communities during the working memory task estimated by modularity maximization. 71

4.10 Modularity scores of overlapping communities for rest and both tasks. The histograms depict the whole-brain modularity scores across 5,000 iterations (each modularity score was obtained by summing modularity scores across communities). 72
4.11 Bottleneck bridges. Bridgeness scores for each region and condition (top: resting state, middle: working memory, bottom: emotion). Colors indicate the percentile of the ROI's median score across 5000 iterations (for example: regions colored red had bridgeness scores around the 90th percentile or above). Black contours indicate regions discussed in the text. Bottom row: Horizontal slices at illustrating strong bridges in the anterior insula. 74
4.12 Hub bridges. Bridgeness scores for each region and condition (top: resting state, middle: working memory, bottom: emotion). Colors indicate the percentile of the ROI's median score across 5000 iterations (for example: regions colored red had bridgeness scores around the 90th percentile or above). Black contours indicate regions discussed in the text. 74
4.13 Membership values for all ROIs with nonzero values reordered in ascending fashion during the working memory task. Green: median; blue: 2.5th percentile; red: 97.5th percentile. 75
4.14 Membership values for all ROIs with nonzero values reordered in ascending fashion during the emotion task. Green: median; blue: 2.5th percentile; red: 97.5th percentile. 76

5.1 **Stimuli paradigm.** To create anxious states, over a period of 60 seconds, two circles with different colors moved around on the screen with some degree of randomness. When they collided with each other, an unpleasant mild electric shock was delivered. 84
5.2 **Intersubject correlation (ISC) and network analysis (ISN). A1)** In ISC the correlation between the same region across different participants' brains is calculated. **A2)** To calculate ISC, for each voxel or region of interest (ROI), the time series "left out" subject (S_1) extracted, and correlated with the average time series of all other subjects (S_2, S_N), for the same voxel/ROI. **A3)** To calculate the group level ISC, the results from A2 across all subjects are averaged. This creates a vector that contains the correlation of every voxel with itself (across participants). **B1)** In ISN the correlation between all pairs of ROIs across different brains is calculated. **B2)** To calculate ISN, for each voxel or ROI, the time series of a "left out" subjected (S_1) is extracted and its correlation with the average time series across other subjects ($S_2, , S_N$) is calculated. **B3)** The group level ISN is a $N_{ROIs} \times N_{ROIs}$ matrix, which shows the average of ISN from all subjects. Note that the vector in A3 corresponds to the diagonal of the matrix in B3, illustrating that intersubject networks provide a richer characterization of time series relationships. 85

5.3 **Procedure for data concatenation used in evaluating intersubject network dynamics. Procedure for data concatenation used in evaluating intersubject network dynamics.** Data point y for each time point t was used to compute intersubject networks for approach and retreat, separately. Data point indexes: ROI, condition segment (two approach data segments are illustrated), and time within segment. Briefly, time was used to "slice" and "concatenate" through the ROI time series. Thus, we generated a data time series at t=0 by concatenating all of the t=0 sample (across same-condition segments across blocks and runs), did the same for t=1, and so on. To account for the hemodynamic delay, we discarded the first 5 seconds of each segment (gray part). The resulting data per ROI, time point, and condition, was then investigated in terms of dynamic properties. Time series data were simulated for illustration. 100

5.4 **Intersubject group networks. A)** Intersubject network (ISN) during threat approach periods. ROI order corresponds that that of Table 1. **B-D)** Average ISN values for each network at approach (B), retreat (C), and approach minus retreat (D). The dark rectangles surround all of the blocks with significant values (p ¡ 0.01); note that each of the 9 blocks was treated as a unity (the outline extended over multiple of them for diagramming convenience). The color bars indicate differences in intersubject correlation. 105

5.5 **Intersubject functional connections between the salience network and subcortical regions. A)** Amygdala La R; **B)** PAG R; **C)** Habenula R; **D)** Cerebellum Crus L. **E)** BNST R. L/R: left, right; Amygdala La: lateral amygdala, PAG: periaqueductal gray, BNST: Bed nucleus of the stria terminalis. For statistical tests, the entire subcortical region was treated as a unit and cohesion between the region and the salience network was tested. Regions shown with green bar plots were significant (p¡0.05). 107

5.6 **Temporal evolution of cohesion during approach and retreat segments.** Within- and between-network cohesion during approach and retreat for the salience, executive, and task-negative networks. For example, as the circles approach each other, the cohesion within the salience network increased; when the circles retreated, the cohesion within the salience network gradually decreased. The orange line shows cohesion values for approach at different times (with a 90% confidence band); the cyan line shows cohesion values for retreat at different times (90% confidence band). The red/blue line shows the least-squares linear fit to cohesion values during approach/retreat; solid lines indicate fits that were statistically significant (p ¡ 0.05). Time is in seconds (TR = 1 sec); the y-axis shows cohesion (summed degree). The red star indicates that the slope difference was statistically significant (statistical values provided at the bottom left). 109

5.7 **Temporal evolution of cohesion between subcortical regions and the salience network.** Although cohesion did not increase robustly during approach periods, cohesion decreased as a function of time during retreat for the left BNST, right habenula, and right PAG. Conventions as in Figure 5.5. 110

A.1 Membership values for all ROIs reordered in ascending fashion during rest. The elbow was defined as the point at maximal distance from the line passing from the first to the last points of the curve (green line). OC1-OC6: overlapping communities during rest (see Figure 3.6). 124
A.2 Overlapping community organization during rest at the individual level for sample participants. Each row depicts the six overlapping communities extracted with the mixed-membership model per subject (sOC). Membership values are thresholded at 0.1 for illustration. . . . 125
A.3 Histograms of membership values at the individual level (N=94). The results illustrate overlapping community structure at this level of analysis. Membership values were thresholded at 0.1 (thus, the lowest bins are always empty). Subject numbering in the figure does not correspond to the one in Figure A.2. 126
A.4 Figure A.3 Continued 127
A.5 Figure A.3 Continued 128
A.6 Figure A.3 Continued 129

B.1 Overlapping communities detected with non-ICA-fixed preprocessing (see text). The color of cortical and subcortical regions reflects the membership value of each region to each community. Membership values are thresholded at 0.1 for illustration. 135
B.2 Cosine similarity between overlapping communities at rest (non-ICA-fixed) and during the working memory task (WM). The matrix displays the cosine similarity between rest and WM membership values. 136
B.3 Cosine similarity between overlapping communities at rest (non-ICA-fixed) and during the emotion task (EM). The matrix displays the cosine similarity between rest and EM membership values. 136
B.4 Relationship between community size and relative bridge type (RBT). For different community sizes, we determined the RTB index. In these plots, the percentile threshold for high bridgeness score was 95%. For each community size, 5000 bootstrapping samples were computed, and the RBT was computed for each; the mean and the 25-75% ranges are shown. The linear fits are for illustration only. . . 137

B.5 Summary results for the relationship between community size and relative bridge type (RBT). The same procedure as in Figure B.4 was employed, but here each matrix entry summarizes the relationship between relative bridge type and community size for a given community and bridgeness threshold; thus, for each cell, community size was varied from 50 to 250 (and in each case, 5000 bootstrapping samples were employed). The relationship was assessed via spearmans rank correlation between RBT and community size. 137
B.6 Modularity scores of overlapping communities during rest ($k = 6$), working memory ($k = 5$) and emotion ($k = 4$). The histograms depict the whole-brain modularity scores across 2,500 iterations (each modularity score was obtained by summing modularity scores across communities). 138

Chapter 1: Introduction

1.1 The Brain: A large-scale complex system

The human brain is one of the most complex living structures which has enabled human beings to achieve paradigm-shifting breakthroughs, from going to space to composing masterpieces of art and music. The physical structure of the brain (its anatomy) consists of billions of neurons and their supportive counterparts, the glial cells (Purves et al., 2008). The neurons are the basic units of brain structure and function. Neurons are highly connected in an organized manner and communicate with each other via electrical and chemical signals (action potentials, synaptic transmission and receptor potentials). On top of this intense physical structure, the brain functions as a complex dynamic system. It continuously receives internal and external stimuli from multiple sources in real time, and uses its neural circuits to integrate new information with previous information, and perform functions (behavior). The interconnected neurons form different types of neural circuits which mediate different functions such as sensory information processing and perception, language, memory, emotion, social cognition, executive control, and reasoning (Purves et al., 2008). To understand the brain, it is necessary to characterize the spatio-temporal interactions of neural elements. This, in turn, motivates investigating the mapping between structure and function of the brain from an integrative and large-scale

network-based perspective (Pessoa, 2014).

Understanding the neural basis of how the brain performs a function, and its relationship to behavior, will enable us to find new ways to treat, cure, and prevent brain disorders, neurodegenerative diseases and mental illnesses. Thus, understanding the brain organization has significant benefits for the society and public policy decision makings in different areas such as health and education.

1.2 Thesis outline

In this research, the spatial and temporal aspects of large-scale functional networks of the human brain during "rest" and "task" conditions are investigated using functional Magnetic Resonance Imaging (fMRI). A review of large-scale brain networks and principles of fMRI are provided in Chapter 2. This dissertation consists of three studies.

In the first study (Chapter 3), we characterize the overlapping structure of large-scale functional brain networks based on "rest" human functional MRI data. To estimate the community structure of the brain, a large body of previous studies has employed techniques which make the assumption that brain regions should be partitioned into disjoint sets of clusters. Though this type of functional brain characterization reveals important information about the brain organization, it cannot capture the flexible and task dependent mapping between brain regions and their functions. Thus, the overlapping structure of functional brain network has been poorly understood. To address this important issue, we studied large-scale brain networks in "rest" condition by employing a mixed-membership Bayesian model

(Gopalan and Blei, 2013), that allows each brain region to belong to multiple communities simultaneously with varying membership strengths. This approach allowed us to compare the structure of disjoint and overlapping communities of the brain. A central finding of our study was that the overlapping communities not only exhibited the general spatial features known from disjoint community estimation of functional intrinsic brain networks, it revealed that at least 60% of regions in each community belong to multiple communities, which illustrates how much information is lost when large-scale networks are treated as disjoint compared to when overlapping structure is characterized. Furthermore, we investigated if regions play roles in multiple communities as a result of the regions participating in different "region assemblies" depending on task demands. We found that the regions' functional diversity (to what extent each region engage in several tasks) is positively related to its membership diversity (to what extent each region participates in several communities) in the absence of a task. This study was presented at the *Pattern Recognition in NeuroImaging (PRNI2015)* conference (Najafi et al., 2015), and has been published in *NeuroImage* Journal (Najafi et al., 2016).

While the large-scale structure of brain networks at rest has been studied extensively, less is known about the large-scale functional network of the brain during task performance. To address this important gap, in Chapter 4 we characterize how overlapping network structure during rest may be altered by task execution (specifically by working memory and emotion tasks). Our findings revealed the existence of dense community overlap during task performance that was not limited to specific regions. Also, the findings revealed important differences between community orga-

nization during rest and during specific task states. In addition to comparing the overlapping communities at rest and task, we extended the modularity definition for mixed-membership algorithms to characterize the degree of non-modularity (or communicability) in brain networks. The results illustrated that the brain communities are more communicating (less-modular) during performing tasks than rest.

Furthermore, by using overlapping communities, we could define hub sub-types by characterizing each node's *bridgeness*, namely, the ability to participate with multiple communities simultaneously and "bridge" them together. We combined the mixed-membership model and network cohesion properties to develop a new measure of the bridgeness to extend our understanding of the universal roles of regions and identify key information processing nodes in the brain. Estimating the bridges in the brain allowed us to understand which regions involve in spreading the signals across multiple communities, thereby performing important roles in distributed processing. This study has been published in *NeuroImage* Journal (Najafi et al., 2016).

Taken together, in Chapters 3 and 4, we suggest that dense overlapping communities are well suited to capture the flexible and task-dependent mapping between brain regions and their functions.

Finally, in Chapter 5 we studied intersubject brain network organization during dynamic anxious anticipation. Specifically, we wished to investigate the evolution of intersubject network properties as threat level changed dynamically. Understanding the neural basis of anxious anticipation is important because aberrant responses to uncertain future negative events are believed to be central to anxiety disorders. Although previous studies have investigated how brain responses

are sensitive to threat proximity, little is known about how patterns of response co-activation change during dynamic manipulations of threat. To address these important gaps in the literature, we studied the dynamics of emotional processing at the level of large-scale brain networks by devising a manipulation in which threat was dynamically modulated during fMRI scanning. Threat was dynamically modulated as two circles slowly meandered on the screen; if they touched, an unpleasant shock was delivered. We extended the intersubject correlation analysis (Hasson et al., 2004) technique to the network level, named as intersubject networks. The intersubject network paradigm allowed us to analyze the synchronization of mental states that are not simply explained by common evoked responses to perceptual features or cognitive demands in network level, and has the advantage of increasing the signal-to-noise ratio. To determine how network properties changed during periods of approach (circles moving closer) and periods of retreat (circles moving apart), we developed the dynamic intersubject network method, and calculated the cohesion within and between the networks. A central finding of our study was that dynamic threat altered network cohesion across the salience, executive, and task-negative networks, as well as in subcortical regions. Functional connections between multiple subcortical regions and the salience network also increased during approach vs. retreat. Together, our findings unraveled dynamic properties of large-scale networks across participants while threat levels varied continuously, and demonstrate the potential of characterizing emotional processing at the level of distributed networks. The results of this chapter was presented at the annual meeting of the *Society for Neurosience* (Najafi and Pessoa, 2016) and is under review for journal publication;

its preprint is available on *bioRXiv* (Najafi et al., 2017).

Chapter 2: Background

2.1 Brain Networks

The recent advancement of neuroimaging techniques, along with computational and theoretical tools and frameworks, has opened new doors to analyze the functional and structural organization of the brain from an integrative and network-based perspective. A network (or graph), is a mathematical representation of a collection of interconnected objects. A network is formulated by $G = (V, E)$, in which V is the set of nodes (vertices) that represent the objects. E is the set of links (edges) comprised of two-element subsets of V that illustrate the connections between the objects. The organization of the brain has been characterized via three different types of networks: 1) structural networks, 2) functional networks, and 3) effective networks. The focus of this dissertation is on functional networks. However, for completeness we will briefly review all types of large-scale brain networks in this section.

2.1.1 Structural network

The structural network of the brain is comprised of spatial nodes and their physical connections. The nodes are neural elements, the definition of which depends on the scale of size which ranges from an individual neuron to entire brain regions. The

same can be said for the links, which are typically delineated as axon projections or neural tracts. Due to the invasive nature of extracting the anatomical connections through techniques such as tract tracing, currently it is not possible to employ the cell-scale methods on a large sample size of the human brain. Therefore, the human structural networks have been studied using diffusion imaging and tractography, which extracts the physical connections based on measurements of the diffusion anisotropy of water or other small molecules within biological tissue. The structural network analyses showed that the structural connectivity of the human brain can be relatively stable on a short time scale of seconds to hours, though changes have been observed at longer time-scales as a result of neuroplasticity. Additionally, the results have shown that the probability of structural connections between regions that are spatially closer is higher than the probability of connections between spatially remote regions. This finding suggests that the wiring in the brain, like the connections in other complex systems, is costly, and the cost of wiring has been an important feature in the evolution of brain networks (Bullmore and Sporns, 2009; Sporns, 2013).

Though understanding the structural connectivity provides the information about the physical system behind the functional interactions between brain regions, studying solely the structural network is not enough to understand how brain regions contribute to brain function (or behavior). Because strong functional connectivity between two brain regions does not necessarily emerge from strong anatomical connections,three different relationships between structural connectivity and functional connectivity exist (Pessoa, 2014): 1) Two functionally connected brain regions are

directly connected via structure, 2) Two brain regions are functionally connected while there is no direct structural connection between them, and 3) Two structurally connected regions do not have functional connectivity. Furthermore, the functional connectivity between two brain regions can be changed based on the task and context without modification of structural connections (Pessoa, 2014). Thus, studying the functional connectivity of the brain regions is important to understanding the brain organization.

2.1.2 Functional network

The functional networks of the brain are constructed based on the neural elements, as nodes, and their "functional" connections, as edges. Functional connectivity characterizes how brain regions interact to generate a function. Friston (Friston et al., 1997) defined the brain's functional connectivity as "temporal correlation between spatially remote neurophysiological events." Functional connectivity is computed by calculating an association metric that characterizes the functional interactions and statistical dependency between neurophysiological time-series of brain regions such as Pearson's correlation, mutual information, or spectral coherency. The neurophysiological time-series (which are based on the neural elements activities or their hemodynamic responses) can be acquired via invasive or non-invasive techniques such as multielectrode array (MEA), electroencephalography (EEG), magnetoencephalography (MEG), or functional Magnetic Resonance Imaging (fMRI). The functional connectivity of the brain is highly time, condition, and task dependent. For example, the functional connectivity between brain regions is altered by per-

forming emotion and motivation tasks (Kinnison et al., 2012; Padmala and Pessoa, 2011; Pessoa et al., 2002).

Though many studies have investigated the functional connectivity from region-centric perspective (a few studies are network-based) during task performance in the past few years, there has been an increasing interest in finding the whole-brain "intrinsic" functional connectivity. "Intrinsic" functional connectivity refers to the connectivity between brain regions during so-called "resting-state", in which the subject is in passive rest or doing a minimal task such as looking at a fixed point. Though there is a debate about whether to consider the resting-state as a "standard" brain state, or simply as another type of task (for further discussions see Chapter 3; Buckner et al. 2013; Cole et al. 2013), the results from the intrinsic functional connectivity, especially from fMRI studies, illustrated functionally coupled regions across the brain whose functional relations were less known. Thus, resting-state fMRI functional connectivity can be a powerful technique to understand the organization of the brain (see next section and Chapter 3).

In the past decade, despite the dynamic properties of functional connectivity, most of the fMRI studies have studied it in a static way, that is the functional connections between brain regions are implicitly assumed to be constant in time. By growing evidence about importance of studying dynamic functional connectivity in understanding brain organization, several strategies such as sliding window techniques (Allen et al., 2012; Chang and Glover, 2010; Rashid et al., 2014), temporal ICA (Liu and Duyn, 2013), and spatio-temporal regularization (Karahanoğlu and Van De Ville, 2015) have been used to study dynamic structure of functional con-

nectivity. Having a non-stationary and noisy data, such as fMRI, which is highly sensitive to participants' motion, has limited the employed techniques and interpretability of the dynamic results. In a recent paper (Pessoa and Najafi, 2015), we suggested more advanced techniques such as *multivariate dynamic Bayesian covariance estimation* method (Wu et al., 2013) can be useful to study dynamics of functional networks using fMRI data. We will address this issue by a novel point of view, which is studying the dynamics of functional networks by using intersubject networks (for details see Chapter 5).

2.1.3 Effective network

Functional connectivity lets us know how brain regions are statistically dependent and associated, but it does not talk about the causal functional relationship of the brain regions. In other words, it does not talk about the direction of information flow. For example, in a three node network if A $\rightarrow$ B $\rightarrow$ C, the time-series of the nodes are correlated and functional connectivity models them as a triangle network. In cases where the correlation between A and C are caused because of B, the correlation methods such as partial correlation can regress out the B effect, and do not consider it as a triangle correlation, but still cannot understand the model and direction between the three nodes. Thus, "functional" connectivity is contrasted from "effective" connectivity in the literature (Smith et al., 2012).

Effective connectivity estimates the directed causal effect between brain regions, i.e effective connectivity exhibits the influence of one brain region on another one. To estimate the directionality, several methods such as multivariate auto-

regression, Bayesian modeling, and higher order statistics have been suggested. For example, multivariate auto-regression models are used to determine whether a region's timeseries is a temporally shifted version of another brain region's timeseries. In this case, the correlation between two regions will be from the one with temporal precedence over the other one. Another example of methods is to look at the probability of A given B, and probability of B given A to estimate the direction of causality (Smith et al., 2012).

2.1.4 Brain Communities

Modern network science has brought many advancements in studying a variety of complex systems. One of the main features is detecting the communities (also named as partitions, clusters, networks, etc.) in the networks. Though there is no agreed-upon or rigorous definition for a community, the main idea is usually based on the concept that a cluster or community is a set of nodes which are "close" to each other according to a distance or similarity measure. Plenty of methods have been introduced to detect the communities in a network: from modularity and centroid-based clustering to clique-based methods and statistical inference. The main idea behind a large number of these methods, such as modularity (Newman, 2006), is based on maximizing the number of within-community links to the between-community links.

The majority of brain network studies have focused on identifying disjoint communities in the brain at rest condition. A central finding is that, at rest, brain regions can be grouped into a relatively small number of stable communities (Balen-

zuela et al., 2010; Power et al., 2012; Yeo et al., 2011). For example, Yeo et al. (2011) described a seven-community parcellation of cortical areas (Figure 2.1) at resting-state by applying k-means on whole-brain functional connectivity based on a very large sample of participants N=1000 (a set of 500 subjects was used as discovery set, and the results was reproduced in the replication set which consisted of the other 500 subjects). Based on anatomical and functional considerations, the communities were labeled as "somatomotor", "visual", "frontoparietal", "default", "ventral attention", "dorsal attention" and "limbic". The first two communities are more related to the sensory regions of the brain (motor movements, and visual). The "default-mode" or "task-negative" network refers to the regions that their activity during passive states (such as rest) increase, while the activity of regions in the "frontoparietal", "ventral", and "dorsal" attentions, may increase by performing demanding tasks (changes in their activation depends on the type of the task). But, as also mentioned by the authors, the focus on seven networks should not be taken to mean other parcellation schemes are not correct, and naming the communities does not mean that the functionalities of the regions in each community are only limited to those of their assigned labels.

Though, the disjoint characterization of functional brain reveals important information about how brain regions group together, it cannot capture the flexible and overlapping mapping between brain regions and their functions. To study large-scale functional brain networks, other techniques such as ICA (e.g., Smith et al. (2012)) and variants of Principal Component Analysis, which allow the components to overlap spatially. To decompose the adjacency matrix, a frequent assumption is that

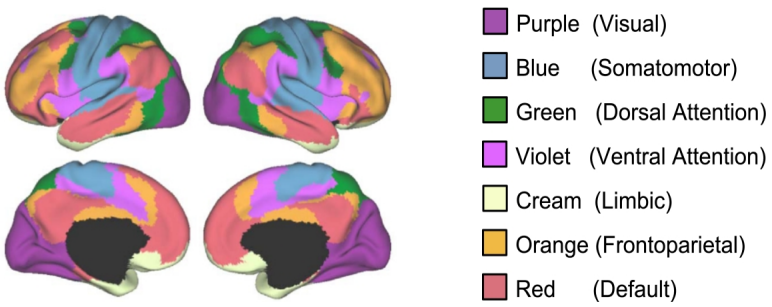

Figure 2.1: The 7-network cortical parcellation of resting-state fMRI data by Yeo et al. (2011). Based on anatomical and functional considerations, the communities were labeled as "Visual", "Somatomotor", "Ventral attention", "Dorsal attention" , "Limbic", "Frontoparietal", and "Default". The figure is adopted from Yeo et al. (2011).

the communities are orthogonal. While there are methods which relax this orthogonality constraint, they all make assumptions about the nature of the underlying sources: for example, temporal ICA assumes that each component will be temporally independent; factor rotations make assumptions regarding sparsity, and so on. Methods based on ICA have made important contributions to the understanding of large-scale brain networks (Calhoun et al. 2001; Smith et al. 2009). Nevertheless, their assumptions are not without controversy (e.g., Friston 1998). In particular, application of temporal ICA is challenging in the context of functional MRI data because it requires a large number of samples to function well, and those are typically not available with standard slow sampling rates ($\sim$2s; though faster sampling with newer pulse sequences reduces this problem).Therefore, the overlapping structure of functional brain network is still poorly understood. We will address these issues in Chapters 3 and 4.

2.2 functional Magnetic Resonance Imaging

fMRI is a noninvasive neuroimaging technique using MRI to measure changes in the brain's functional activity. fMRI, however, does not construct images of neural activity itself; instead, it creates images of physiological activity correlated with neuronal activity.

2.2.1 Principles of fMRI

The main metabolic sources of energy for the neurons are glucose and oxygen, which are provided by the vascular system. Oxygen is carried to the neurons by oxygenated hemoglobin (HB) molecules. Once it has deposited its oxygen, hemoglobin becomes deoxygenated hemoglobin (dHB). As the neurons become active, the vascular system must provide more glucose and oxygen to satisfy their energy requirements. Thus, the rate of conversion of HB to dHB will increase in that area. In 1936, Linus Pauling and Charles Coryell discovered that the magnetic moments of HB and dHB are different as a result of their electrical properties. HB has no free electron, therefore it is diamagentic. dHB has two unpaired electrons, and therefore has a significant magnetic moment. In 1990, Ogawa and colleagues proposed that the difference in the magnetic moments of HB and dHB can be used to create the blood-oxygenation-level dependent (BOLD) contrast images using MRI. The changes in rate of conversion of HB to dHB because of neurons activation, can be acquired by MRI scanners, as a slight increase in MR signal which leads to a brighter region in fMRI images (Huettel et al., 2004).

2.2.2 MRI scanner and BOLD Imaging

The functional MRI data (like structural MRI data) is acquired by MRI scanners which have three main components: 1) a superconducting magnet to generate the main static magnetic field, 2) radio-frequency coils to collect the MR signal, and 3) gradient coils to provide spatial information of MR signal.

When a body is placed within the static magnet, its atomic nuclei (mainly hydrogen) are aligned with the static magnetic field. Nuclei have intrinsic spin, and this spin causes a precession (imagine a tilted top) about the magnetization direction. The frequency of this precession is known as the Larmor frequency and it is proportional to the strength of the static field and is dependent on the type of nuclei. Increasing the angle of the precession places the nuclei in a higher energy state (excited). This can be accomplished by applying a magnetic force at the Larmor frequency, which happens to be in the range of radio electromagnetic radiation. When the nuclei relax back to the un-tilted state (equilibrium), they produce radio waves with amplitude related to the degree of the induced tilt. This is the MR signal. Radio-frequency coils are needed to collect MR signal by generating and receiving magnetic pulses, and gradient coils are needed to save the spatial information of MR signal. The gradient coils pulse sequence frequency varies based on the modality of imaging (structral, functional, diffusion,etc.). In structural imaging, the density of water molecules which differs based on tissue type is measured, whereas in the functional MRI the BOLD signal will be measured (Huettel et al., 2004).

In comparison with other neuroimaging methods, such as EEG or MEG, fMRI

provides a lower temporal resolution (in range of hundreds of milliseconds to a few seconds), but higher spatial resolution (in range of a few millimeters).

2.2.3 MRI data preprocessing pipeline

The meaningful neural-related variability in fMRI measurements, like many other neuroimaging techniques, is obscured by non-neural-related spatio-temporal variability, or noise. This unwanted variability can be caused by different factors, such as changes in the temperature of scanner to physiological effects like brain pulsation due to heart beat, and head motion. The motion artefacts can be more problematic, because it can have temporal correlation with the experiment. To eliminate the effect of non-neural variability in fMRI, the fMRI data should go through a series of preprocessing steps.

As the first step of preprocessing, the images are visually inspected to prevent and diagnose data problems. Many of the common (large) artefacts are immediately distinguishable. If the data is corrupted significantly by the scanner or physiological noise, the data may be excluded from further analysis. The visual quality assurance is not enough, and quantitative methods are applied to the data to correct for possible temporal and spatial distortions.

Temporal and spatial corrections are needed to compensate for the difference in the timing and spatial position of the slice acquisition. For example, if the slices are acquired in an interleaved format (in which all odd numbered slices are acquired followed by the even slices), the activity in adjacent regions of the brain are acquired with non-adjacent time points. This can make two regions with identical

hemodynamic response appear less correlated. The slice time correction can be done by different methods such as interpolation which uses the information from nearby time points at each time period (TR) to estimate the amplitude of BOLD signal for each slice.

The spatial error can be caused by the inhomogeneity of the static field or head motion. The spatial distortions emerged from inhomogeneity of static field can be corrected by acquiring field maps to estimate the field and distortion amount at each voxel. Head motion, however, is unavoidable and is the main cause of spatial distortions. Small movement can be corrected by realigning and transforming the image volumes so that each voxel position matches its position in a reference image. For larger motion (more than half to 1 millimeter), it has been suggested that the volumes be excluded from further analysis (Power et al., 2012). Motion correction alone cannot correct the motion's statistical effect on the BOLD signals. Therefore, motion parameters, and even their first temporal derivatives, are regressed out from the BOLD signal. Furthermore, it has been suggested that if regressing out the motion parameters does not minimize the temporal effects of motion on BOLD signal at specific time points, those time points may be removed to avoid spurious patterns of co-activation because of the head motion (Power et al., 2012).

To improve the spatial localization of activation, each subject's functional data is registered to its structural image. This is accomplished via linear transformations such as rigid-body or affine, or by using tissue boundary information. The registered images must then be registered to a structural template to allow for inter-subject activity localization comparison as well as comparison to other studies.

Functional data resolution can be improved by applying proper spatial and temporal filters. Spatial smoothing, which is usually done via applying a low-pass spatial Gaussian filter, increases the signal-to-noise ratio (SNR) by combining neural similarity between adjacent regions. Increasing SNR improves the validity of statistical analysis. The temporal filtering can improve the functional data by suppressing non-neural frequency components, such as scanner's drift or those introduced by heart rate and respiration.

Chapter 3: Overlapping communities reveal rich structure in large-scale brain networks during rest

3.1 Introduction

Network analysis of human neuroimaging data has contributed to a view of brain function in which groups of brain regions participate in functions rather than brain function relying on just regions operating in isolation (for influential early ideas, see Mesulam et al. (1981); Rakic et al. (1986)). Functional MRI data during the so-called "resting state" has been extensively investigated in order to characterize network structure. A central finding is that, at rest, brain regions can be grouped into a relatively small number of stable communities, also called clusters or sub-networks. For example, Yeo and colleagues (Yeo et al., 2011) described a seven-community parcellation of cortical areas based on a large sample of participants. Based on anatomical and functional considerations, the communities were labelled as "visual," "frontoparietal," "default," and so on (Figure 2.1).

Much of the work employing modern network methods to study brain community structure and other network measures makes the assumption that each node (that is, brain region) belongs to a single community thus, the overall network is partitioned into disjoint sets of clusters. However, the importance of understanding and characterizing overlapping structure has been discussed for some time in many

disciplines, including sociology (Wasserman and Faust, 1994) and biology (Gavin et al., 2002); for example, biologists exploring protein interactions have found that a substantial fraction of proteins interact with several communities at the same time (Gavin et al., 2002). As nicely summarized by Palla and colleagues (Palla et al. (2005), p. 814): "actual networks are made of interwoven sets of overlapping communities." Indeed, there is increasing realization that brain regions may belong to several brain communities simultaneously (Cocchi et al. (2013); Cole et al. (2013); Pessoa (2014); Smith et al. (2012); Yeo et al. (2014); see also Hilgetag et al. (1996); Mesulam (1998)). It is still unclear, however, whether network overlap is sparse or dense (Figure 3.1). In the former case, some key regions work as hubs that participate flexibly in multiple communities. In this scenario, network overlap is relatively limited and may be a property mainly of specific parts of the brain (say, prefrontal cortex; see Miller and Cohen (2001)). In the latter case, network overlap is extensive and many (possibly most) brain regions participate in multiple communities.

The first goal of this chapter is to characterize the overlapping structure of brain communities during rest by using a state-of-the-art mixed-membership algorithm (Gopalan and Blei (2013); see also Airoldi et al. (2014, 2008); Lancichinetti et al. (2009)). In particular, our goal was to quantify how much "information" is lost when large-scale networks are treated as disjoint compared to when overlapping structure is characterized. Standard, disjoint clustering assigns membership values of 0 or 1 (all or none), and in so doing dichotomizes measures that may be informative. In contrast, when overlapping structure is determined, a node's participation

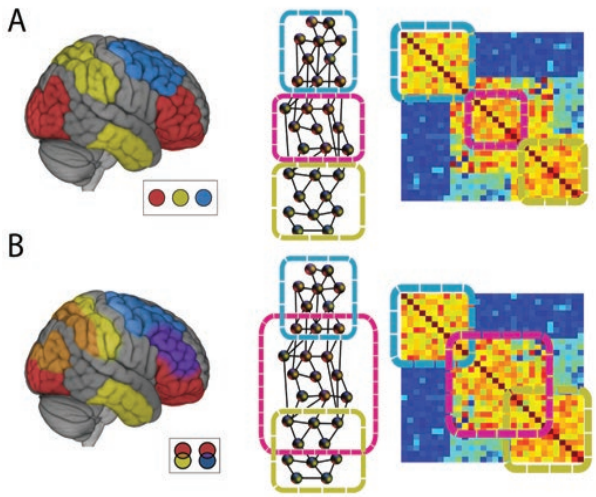

Figure 3.1: Community organization in the brain. (A) Standard communities are disjoint (inset: colors indicate separate communities), as illustrated via three types of representation. The representation on the right corresponds to a schematic correlation matrix. (B) Overlapping communities are interdigitated, such that brain regions belong to multiple communities simultaneously (inset: community overlap shown on the brain indicated by intersection colors).

is assigned across all communities, though with varying strengths; the strengths are summarized by the membership values (Figure 3.2). Specifically, in the framework adopted, each node has a probability-like membership value associated with each of the existing communities. This community membership vector specifies a node's affiliation to all communities considered, with membership values between 0 and 1 (and summing to 1), with entries close to 1 indicating membership to essentially one community, and intermediate values indicating membership to multiple communities.

One reason regions may belong to (or affiliate with) multiple communities is that they may participate in different "region assemblies" depending on task demands. Whereas evidence for regions with "flexible functional connectivity pat-

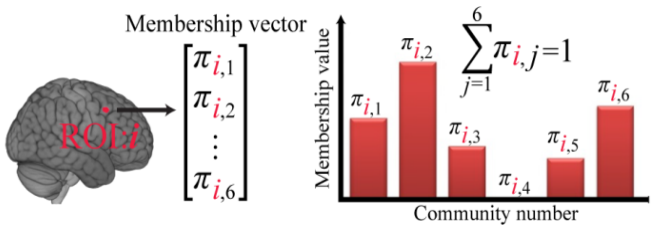

Figure 3.2: Overlapping communities and membership values. Each brain region affiliates with each community with varying strengths that are captured by the membership value. These probability-like values are between 0 and 1 and sum to 1 (for each region of interest, or ROI).

terns" in frontal and parietal cortex has been recently described (Cole et al., 2013), the extent of such multiple-community participation, and its spatial distribution in the brain remains poorly understood. Moreover, the relationship of flexible affiliation and functional diversity (that is, the spectrum of tasks a region may participate in; see Passingham et al. (2002)) is not understood. Therefore, we sought to investigate the relationship between functional diversity and community membership properties. Brain regions differ in terms of their functional diversity, namely, the repertoire of functions that engage them (Passingham et al., 2002). Some regions are engaged by a wide variety of tasks (they have high functional diversity), whereas other regions are more narrowly tuned and are engaged by a limited range of paradigms (they have low functional diversity). Here, we asked the following question: Is functional diversity related to how brain regions affiliate with other regions in the absence of a task? In other words, is a region's functional diversity related to its membership values? To estimate a region's functional diversity, we employed the BrainMap database (Laird et al., 2005), which collates activation results across thousands of publications in the literature and organizes them in terms

of a task taxonomy spanning perception, action, cognition, and emotion.

3.2 Materials and Methods

3.2.1 Resting-state data

The structural and functional MRI data for this study were obtained from the Human Connectome Project (HCP; Van Essen et al. (2013)) dataset (N = 100) as accessed in June 2014. For completeness, we briefly describe the main aspects of the HCP data (for details, see Glasser et al. (2013); Smith et al. (2013); Van Essen et al. (2013)). Data were acquired on a Siemens Skyra 3T scanner using a standard 32-channel head coil. Functional data were collected using a multiband scanning protocol (multiband factor of 8) that allowed acquisition at higher temporal (TR = 720 ms; TE = 33.1 ms; FA: 52) and spatial resolution (2 mm isotropic voxels in 72 slices; FOV= 208 180 mm). Here, we employed HCP's so-called "ICA-fixed Resting-state" functional data (Salimi-Khorshidi et al., 2014; Smith et al., 2013), which included the following preprocessing steps: correction of spatial distortions, motion correction, functional to structural data alignment, bias field correction, and intensity normalization. Cortical data were mapped to the surface (using the Conte69, 32k standard mesh), and subcortical data were analyzed volumetrically. Surface data were spatially smoothed within the surface with a 2 mm kernel. Finally, an ICA based technique is employed to remove spatiotemporal components that purportedly corresponded to non-neural signals. We removed data from six participants due to scanning artifacts (see section 4.2.1.3 for details)

Experiment design

Each run of resting-state data contained 1200 volumes (14 min and 33 s). Participants were instructed to look at a fixation crosshair for the duration of the run with their eyes open. We employed data from two runs acquired on each participants' first visit. The only preprocessing step applied beyond the HCP ICA-fixed pipeline was to perform temporal filtering (between 0.01 and 0.08 Hz) via the 3dBandpass program of the AFNI package (Cox, 1996).

3.2.2 Regions of Interest

Cortical regions of interest (ROIs) were defined to cover the cortex, yet be small enough to minimize the mixing of signals from adjacent but (potentially) functionally heterogeneous regions. k-means (with "city block" or L1 distance; 1000 iterations; Duda et al. (2012)) was used to cluster each cortical hemisphere into 500 target regions based on their xyz coordinates. Thus, ROIs were defined based on their spatial coordinates on the cortical mesh, not functional data. Because the medial surface of the hemisphere is naturally undefined for parts that are effectively subcortical, clusters with xyz coordinates in those locations (i.e., corresponding to subcortical locations) were not considered. Therefore, a total of 941 ROIs were defined (471 on the right hemisphere and 470 on the left hemisphere). The resulting clusters were further "eroded" so as to avoid ROIs from abutting one another to minimize signals from a given region from contributing to neighboring ROIs (on average, ROIs comprised 28 2x2x2 mm voxels). Subcortical regions of interest were

defined anatomically based on Freesurfer's subcortical segmentation (Fischl, 2012). Each subcortical structure was considered as one region, resulting in a total 19 subcortical regions (9 in each hemisphere and one brainstem ROI), including the cerebellum.

3.2.3 Functional connectivity

Functional connectivity between every pair of regions was calculated using the Pearson correlation of regions' mean time series, resulting in a 960-by-960 connectivity matrix (for each participant). Here, we followed the strategy by Yeo, Buckner and colleagues (Yeo et al., 2011), Power and colleagues (Power et al., 2012), and Cole and colleagues (Cole et al., 2013) and binarized each participant's connectivity matrix by setting the top 10% strongest correlations to 1 and all other entries of the correlation matrix to 0. This strategy has proved successful in producing stable community results. The binarized connectivity matrices were averaged across participants and binarized a second time; again, the top 10% strongest connections were set to 1 and all other values were set to 0. The latter binarization was used so that we could apply the Bayesian mixed-membership model (see below) to partition the correlation matrix, which requires a binary correlation matrix. Note that the two-step binarization used here is not typically employed by other investigators, as some prefer to threshold at the subject level and others at the group level; our approach simply combined subject-level and group-level thresholding (the present approach is nearly indistinguishable from using a single thresholding step at the group level;

difference in only 0.8% of the edges). Note, furthermore, that negative correlations were converted to zero. While this practice is widespread in analysis of brain data, there is also increasing realization of the potential importance of negative weights (Rubinov and Sporns, 2011). Note, however, that our thresholding method would also eliminate all negative weights because these were relatively weak links (over 99% of the negative edges were in the bottom 20th percentile of correlation strengths).

3.2.4 Detecting disjoint communities

Clustering with k-means (Duda et al., 2012) as implemented in Statistics and Machine Learning toolbox in MATLABR2012b (with "city block" or L1 distance; 1000 iterations) was applied to the averaged adjacency matrix to identify k disjoint communities. Although k-means is a very simple clustering method, we employed it here for consistency with previous studies such as Yeo et al. (2011), whose results were based on 1000 participants.

3.2.5 Detecting overlapping communities

The approach to detecting overlapping communities adopted here was originally based on the mixed-membership stochastic blocks model (Airoldi et al., 2008) within the context of stochastic variational inference (Hoffman et al., 2013). The specific algorithm we employed was developed by Gopalan and Blei (Gopalan and Blei (2013); https://github.com/premgopalan/svinet).

3.2.5.1 Mixed-membership algorithm

The model by Gopalan and Blei (2013) can be considered to belong to a subclass of mixed-membership stochastic block models (Airoldi et al., 2008) that assumes assortativity.

To understand the model, let us consider it from a generative standpoint (i.e., as an engine to generate data according to certain distributional assumptions). To generate a network, the model considers all pairs of nodes. An essential component of the model is that it implements assortative clustering: nodes with similar membership vectors are more likely to be connected to each other. To capture assortativity, for each node pair (i, j), the model specifies *community indicator* vectors $\mathbf{z}_{i \to j}$ and $\mathbf{z}_{i \leftarrow j}$. Community vectors have dimensionality $1 \times K$ and have all entries set to zero, except for the entry that equals 1 and indicates which community the node is likely to belong to. If both indicators "point" to the same community (i.e., are 1 for the same entry of the vector), the model connects nodes i and j with higher probability; otherwise they are likely to be unconnected.

Each community is associated with a community strength, β_k , which captures how connected its members are. The probability that two nodes are connected in the model is governed by the similarity of their community membership vectors and the strength of their shared communities (the latter via the parameters β_k). Formally, the model assumes the following generative model to define networks (Algorithm 3.1, Figure 3.3). The notation "$\sim$" stands for "distributed as".

From the complementary standpoint of parameter estimation, given an ob-

Algorithm 3.1 The mixed-membership community detection generative model.

1) For each community k, sample community strength $\beta_k \sim \text{Beta}(\boldsymbol{\eta})$.
2) For each node i, sample community memberships $\boldsymbol{\pi}_i \sim \text{Dirichlet}(\boldsymbol{\alpha})$.
3) For each pair of nodes i and j, where $i < j$:
   - Sample community indicator $\mathbf{z}_{i \to j} \sim \text{Multinomial}(\boldsymbol{\pi}_i)$
   - Sample community indicator $\mathbf{z}_{i \leftarrow j} \sim \text{Multinomial}(\boldsymbol{\pi}_j)$
   - Sample connection $y_{ij} \sim \text{Bernoulli(r)}$, where:

$$r = \begin{cases} \beta_k & \text{if } z_{i \to j} = z_{i \leftarrow j} \text{ (indicators of node } i \text{ and } j \text{ point to same community)} \\ \epsilon & \text{if } z_{i \to j} \neq z_{i \leftarrow j} \text{(indicators of node } i \text{ and } j \text{ point to different communities)} \end{cases}$$

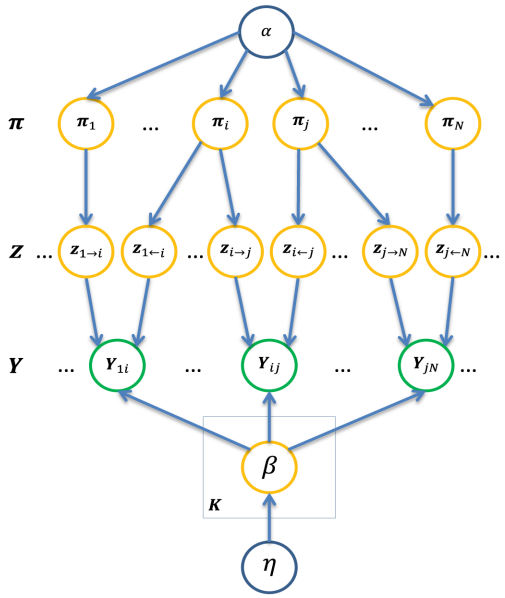

Figure 3.3: Graphical representation of the mixed-membership model. Green circles indicate observed variables; orange circles indicate hidden variables. Box indicates replication over communities. Based on Gopalan and Blei (2013).

served network (the input to the algorithm), the model defines a posterior distribution that gives a decomposition of the nodes into K overlapping communities. Formally, the algorithm computes the posterior distribution, $p(\pi_{1:N}, z, \beta_{1:K} | y, \alpha, \eta)$, so that the community membership vectors can be estimated these are the objects that define the membership values for every node, which define the overlapping

structure of the overall graph. Exact inference of this posterior distribution is computationally intractable, so stochastic variational inference (Hoffman et al., 2013) is used to approximately estimate the posterior in an efficient iterative way (such that estimating the posterior becomes an optimization problem). Importantly, the algorithm is fast and scalable. Indeed, the method has been used to study a network of 575,000 scientific articles (nodes) from the arXiv preprint server and a network of 3,700,000 US patents (nodes) (Gopalan and Blei, 2013).

Because the algorithm is based on a probabilistic generative model, the values of the membership values depend on the initialization of the random "seed." Therefore, we ran the algorithm 100 times with different random initialization seeds. The average membership value across the 100 initializations was considered as the membership value for that community. Of course, across the 100 "runs," the columns of the membership matrix (i.e., the communities) may be "shuffled" (for example, community 1 in one iteration may be community 5 in another iteration). Therefore, prior to averaging across the 100 initializations, a graph matching procedure was developed to reorder the columns (i.e., communities). To do so, the $100K$ different communities were clustered into K groups using k-means clustering (for computational expediency, we used the algorithm's version as implemented in Python's scikit-learn package in http://scikit-learn.org/stable/ which uses Euclidean or L2 distance; 1000 iterations), and ordered accordingly. Note that this step does not determine overlapping community structure, and simply matches community labels. We also tested other matching methods, for instance, using normalized mutual information, and these generated nearly identical results. In essence, the Bayesian

mixed-membership model was fairly robust to seed initialization, leading to stable matching results. Importantly, by calculating the membership matrix based on 100 runs of the algorithm, we are able to mitigate the "degeneracy" problem (Good et al., 2010), namely, the fact that multiple community assignments lead to similar model fits (for further discussion, see Fortunato (2010); Pessoa (2014)).

3.2.5.2 Choosing the number of communities

To determine the number of communities, for each condition, the Bayesian mixed-membership model fit to the data was computed for different numbers of communities, ranging between 2 and 25. To hold the value of k constant across all bootstrapping iterations, we computed the model fit to the data by considering the entire original dataset (N=94). Figure 3.4 depicts the model fit for resting dataset as a function of k. We observed a slight peak at $k = 6$. We note, however, that this does not imply that the the "correct" number of communities at rest in general is 6, as choices of k between 4 and 8 were quite similar; in addition, these values of k are specific to the model employed here, and other models could partition the data optimally with a different number of communities.

3.2.6 Reproducibility and reliability of the results

Instead of splitting the data into "discovery" and "replication" datasets, we applied bootstrapping (Efron, 1979) to our data by randomly resampling subjects (with replacement) for a total of 5000 iterations. Overlapping network analysis was per-

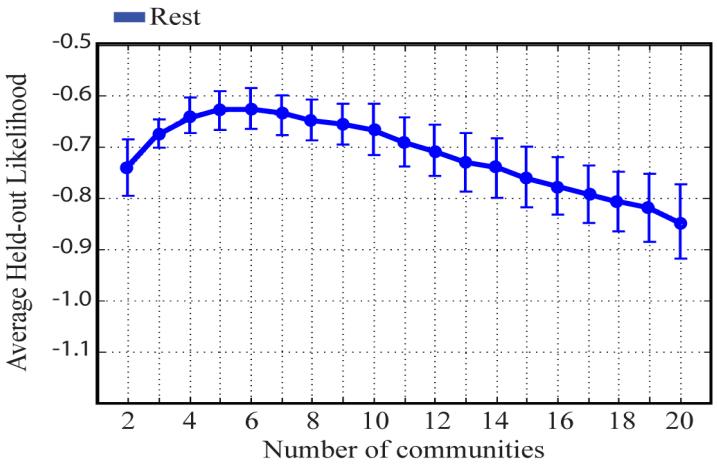

Figure 3.4: Model fit as a function of the number of communities for resting-state dataset investigated. Because the trend continues after $k = 20$, we do not show values $20 < k \leq 25$.

formed independently for each of the 5000 iterations (each sample of 94 participants defined its own functional connectivity matrix).

By estimating overlapping communities for 5000 randomly resampled datasets, we generated 5000 sets of results, which allowed us to estimate median results as well as determine confidence intervals for them. To match community labels across iterations, we employed the same method described above, namely, k-means clustering.

The bootstrapping samples were also employed to determine nodes with membership values that were reliably greater than zero. To do so, to determine its lower bound, we computed the 2.5th percentile of the membership values for a given community (Figure A.1). We then sorted the regions based on their 2.5th percentile membership value, and found the "elbow" of the 2.5th percentile "curve" (the "elbow" was defined as the point at maximal distance from the line passing from the

first to the last point of the curve). Importantly, our results were robust to different choices of threshold.

3.2.6.1 Community similarity

To assess the similarity between the estimated communities from overlapping community detection method and disjoint community detection, we employed cosine similarity (i.e normalized dot product) as follows:

$$\text{simiarilty} = cos(\theta) = \frac{\mathbf{C}_i.\mathbf{C}_j}{\|\mathbf{C}_i\| \, \|\mathbf{C}_j\|} \tag{3.1}$$

where $\mathbf{C}_i$ is the i-th community membership values, and $\mathbf{C}_j$ is the j-th community membership values. The similarity value ranges from 0 to 1: if $\mathbf{C}_i$ and $\mathbf{C}_j$ are identical, the similarity value becomes 1.

3.2.7 Functional diversity

Previously, Anderson et al. (2013) studied the functional diversity of brain regions, namely, the repertoire of tasks that regions are engaged in, by utilizing repositories of human neuroimaging data. Following a similar approach, here we used task activations from the BrainMap database (Laird et al. (2005); http://www.brainmap.org) to estimate functional diversity for each region of interest. We employed activation coordinates from experiments in 47 task domains spanning perception, action, cognition, and emotion in healthy subjects (total of 124,211 activation coordinates, as accessed on November 12, 2014). Activations from clinical populations were not considered. The initial steps of data analysis were performed volumetrically (that

is, voxelwise) because of the meta-analytic nature of the data. Initially, a "domain-frequency map" was computed by tallying the number of times an activation was observed for each voxel (2 mm isotropic voxels), for each of the 47 task domains. The 47 maps were then convolved with a 6-mm cube-shaped kernel, and normalized to have a sum equal to one across domains (per voxel). Thus, the values at each voxel in each domain-frequency map ranged between 0 and 1.

To compute functional diversity, we employed Shannon's entropy measure (Shannon, 2001), which is used across many disciplines (e.g., biology, economics) as a measure of "diversity" (e.g., Magurran (2004)). Formally,

$$H_i = -\sum_{j=1}^{k} p_{ij} \log p_{ij} \tag{3.2}$$

where H_i is the functional diversity of the i-th region, p_{ij} is the frequency of activation in the database in the j-th domain for that region, and $D = 47$ is the number of task domains. Thus, maximal diversity occurs when the region belongs equally to all communities, namely, $p_{ij} = \frac{1}{D}$.

The final voxelwise diversity map was then mapped to the surface (by using HCP's Workbench volume-to-surface-mapping function; Marcus et al. (2011)). For each surface ROI, the p_{ij} values were defined as the average activation frequency across the vertices associated with the ROI. For subcortical regions, no mapping was necessary and p_{ij} values were defined as the average activation frequency across the voxels within the anatomical ROI.

3.2.8 Membership diversity

We computed the diversity of membership values for each ROI by employing Shannon's entropy measure (Shannon, 2001), in a manner that was analogous to how we computed functional diversity. But note that membership diversity measures the extent to which a region has diverse membership values (one value per community; see Figure 3.2); functional diversity (previous section) measures the extent to which a region activates to tasks across domains.

In terms of the entropy equation (equation 3.2), if p_{ij} is the probability that the i-th region belongs to the j-th community, and D is the number of communities, H_i is the membership diversity for region i. Thus, maximal membership diversity occurs when the region belongs equally to all communities. Because each region has membership values $\pi_{i,j}$ (i indexes ROIs, j indexes communities) that take on values between 0 to 1, and membership values for each region sum to 1, we can use the $\pi_{i,j}$ as "probabilities" $p_{i,j}$ in the equation 3.2. For each ROI, the membership diversity was calculated by using estimated overlapping communities from mixed-membership model.

3.2.9 Cortical and Subcortical Visualization

To visualize cortical results we used SUMA (Saad and Reynolds (2012); http://afni.nimh.nih.gov/afni/suma), and to visualize the subcortical volumes we utilized the volume shapes in the CANlab 3dHeadUtility MATLAB toolbox by Wager et al. (2008)); https://github.com/canlab/CanlabCore). To aid visualization,

the right putamen is plotted behind other subcortical regions of the left hemisphere, and the left putamen is plotted behind other subcortical regions of the right hemisphere.

3.3 Results

Analysis of functional MRI data was performed on a sample of 100 unrelated subjects from the Human Connectome Project (Van Essen et al., 2013). We removed data from six participants due to scanning artifacts.

The cortical surface and subcortical volumes were subdivided into 960 regions. Whole-brain functional connectivity was measured for each participant by calculating the Pearson correlation between the time series from every pair of regions. To decompose the functional connectivity matrix into a set of disjoint communities, we employed standard k-means (Materials and Methods), which identifies clusters of regions such that each region belongs to a single community.

For overlapping clustering, we employed a Bayesian mixed-membership model (Gopalan and Blei (2013); Materials and Methods) which identifies sets of regions such that each node can belong to multiple communities. The mixed-membership model assigns a continuous, probability-like membership value for each region of each community (Figure 3.2). Thus, a region belonging to a single community has a membership value of one for that community and membership values of zero elsewhere; a region belonging to multiple communities has intermediate membership values (between 0 and 1) across communities. To be able to assess the reliability of

the results, the non-discarded datasets (N=94) were analyzed by a bootstrapping procedure (Materials and Methods) that allowed us to estimate the median of the parameter estimates reported below, as well as determine confidence intervals for the parameters.

Both disjoint and overlapping clustering algorithms require the specification of the number of communities, K, to be extracted. We report here results based on K equal six, which was determined by evaluating the model fit to the data (Materials and Methods; Figure 3.4).

3.3.1 Comparing disjoint and overlapping communities at rest

The six communities extracted with disjoint clustering (k-means; Figure 3.5) corresponded closely to previously identified large-scale communities (Yeo et al., 2011); but note that unlike most previous studies, we included subcortical regions, too. Here, we avoid using semantic labels for communities (for example, "visual", "frontoparietal", and "default"), and simply refer to the disjoint communities as DC1 through DC6.

The six overlapping communities obtained with the Bayesian mixed-membership model, including membership values for every region, are depicted in Figure 3.6. The correspondence between each pair of disjoint and overlapping communities was measured using cosine similarity.

Four of the disjoint communities exhibited a fairly high degree of similarity with a single overlapping community (Figure 3.7). Two disjoint communities, DC5

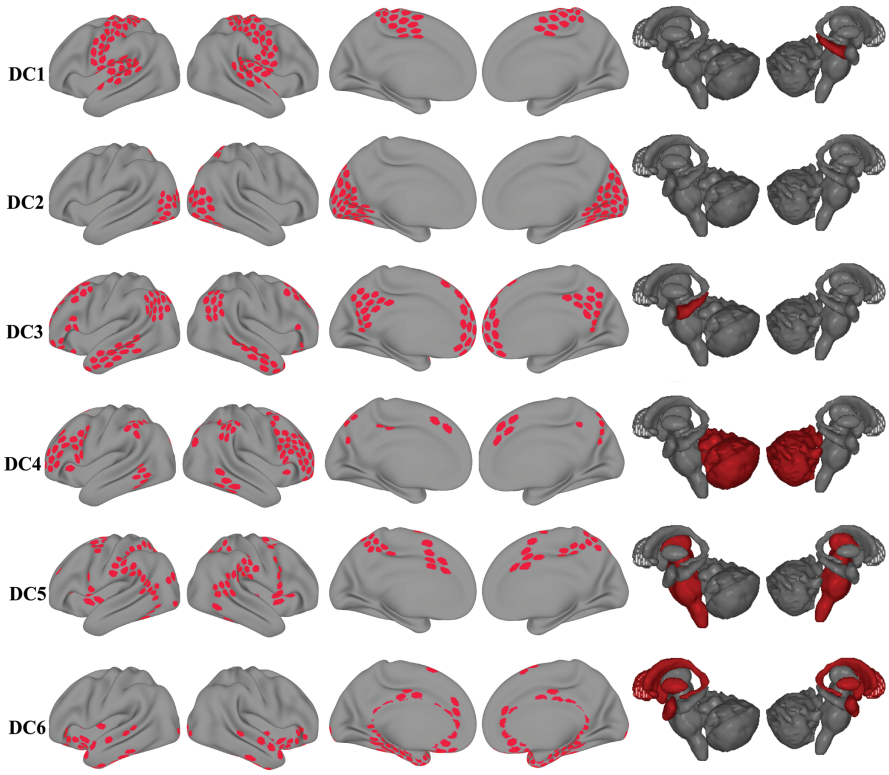

Figure 3.5: Disjoint communities (DC) detected during rest. Each row depicts one of the six disjoint communities extracted with k-means. Cortical and subcortical regions belonging to each of community are colored in red.

and DC6, exhibited a less clear correspondence with a unique overlapping community: DC5 was similar to both OC5 and OC6; DC6 did not correspond strongly to any of the overlapping communities. From the perspective of the overlapping communities, OC5 corresponded most strongly with DC5, which is a network that includes fronto-parietal regions important for attention and executive function. Interestingly, OC5 was least similar to DC3, the network that is commonly denoted as "resting state" or "task negative."

To characterize the finer structure of the overlapping communities, we examined the distribution of membership values for each of the six communities (Figure 3.8). Communities OC1 through OC3 displayed peaks at the largest bin values,

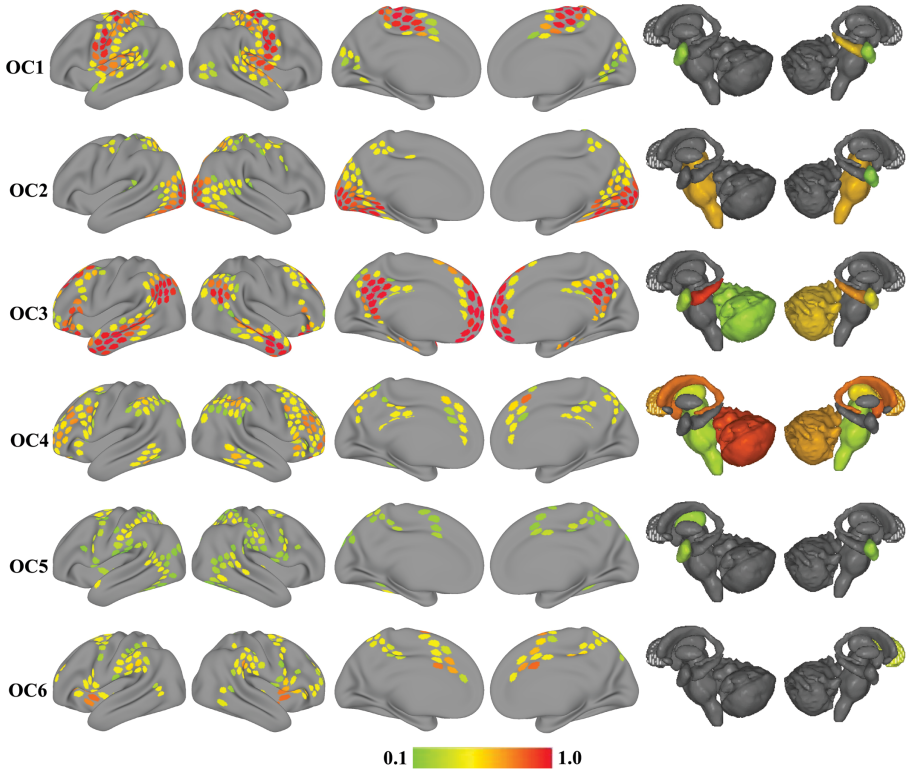

Figure 3.6: Overlapping communities (OC) detected during rest. Each row depicts one of the six overlapping communities extracted with the mixed-membership model. The color of cortical and subcortical regions reflects the median membership value of each region to each community across 5,000 iterations. Membership values are thresholded at 0.1 for illustration.

but also considerable probability mass below the bin with largest membership values. Interestingly, OC3, the community most similar to the standard task-negative network, showed the most skewed distribution, and the only community with a clear peak close to values of 1. In contrast to the increasing pattern of OC1-3, communities OC4 through OC6 showed distinct shapes. OC4 displayed a distribution of values that was relatively uniform (except for the first and last bins); both OC5 and OC6 exhibited negative skew, with OC5 showing particularly strong skew.

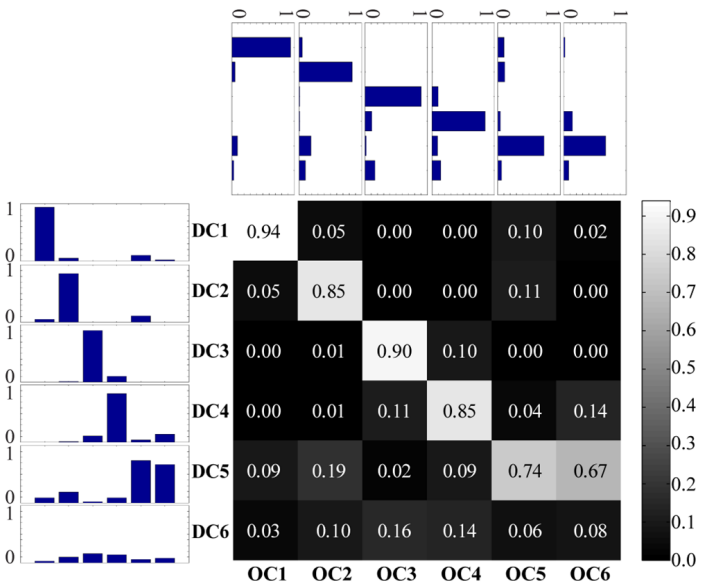

Figure 3.7: Cosine similarity between disjoint (DC1-DC6) and overlapping (OC1-OC6) communities during rest. The matrix displays the median cosine similarity between community pairs across 5,000 iterations. Given that community membership vectors do not contain negative values, the cosine similarity scores range from zero (orthogonal) to one (identical). Side and top insets represent similarity scores as bar plots across rows and columns, respectively.

3.3.2 Relationship between functional diversity and membership diversity

Brain regions differ in terms of their functional diversity, namely, the repertoire of functions they are engaged in. By considering large datasets of neuroimaging studies Anderson et al. (2013) showed that some regions are engaged by a wide variety of tasks (they have high functional diversity), whereas other regions are more narrowly tuned and are engaged by a limited range of paradigms (they have low functional diversity). Here, we asked the following question: Is functional diversity related to how brain regions affiliate with other regions in the absence of a task (resting state)?

To study this question, we defined a new measure for overlapping communities,

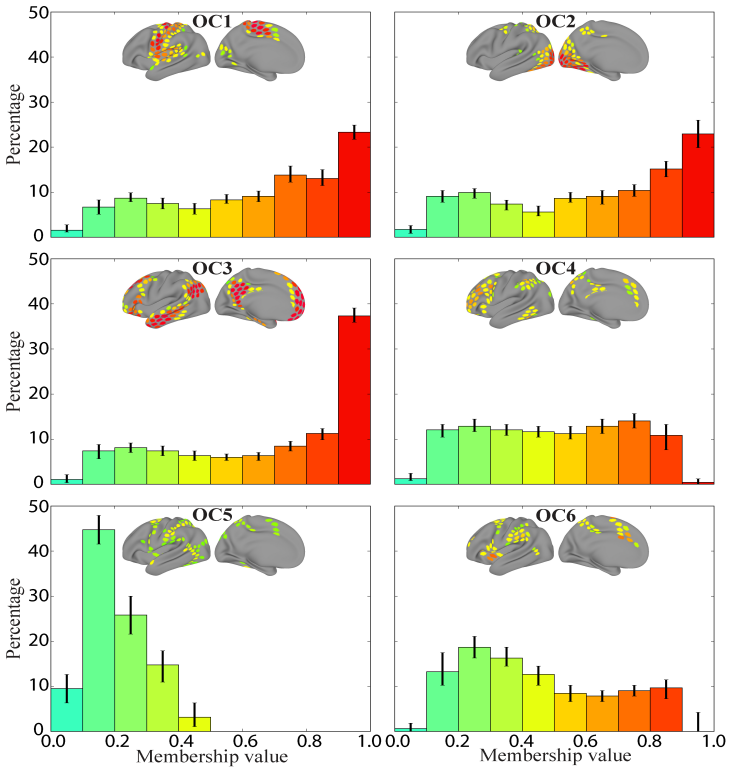

Figure 3.8: Frequency histograms of membership values for each of the six overlapping communities during rest. Each histogram depicts the median value in each bin for that community across 5,000 iterations (error bars show the of 25th-75th percentile range). The colors of the bars correspond to the range of membership values shown in the brain insets.

the membership diversity, which captures the extent to which a node participates in multiple communities. A node's membership diversity was quantified by the Shannon entropy of the membership values (Materials and Methods). Intuitively, maximal diversity occurs when a region participates equally across all communities; naturally, minimal diversity occurs when a region participates in one and only one community. Separately, functional diversity was calculated (see Materials and Methods; Figure 3.9) for each region based on neuroimaging data in the BrainMap database (Laird et al., 2005). Robust regression revealed that functional diversity was positively related to membership diversity, indicating that regions activated by

a wide variety of tasks (that is, functionally diverse) tended to participate in more communities at rest (median slope: 0.27, 95% confidence interval: 0.22-0.32).

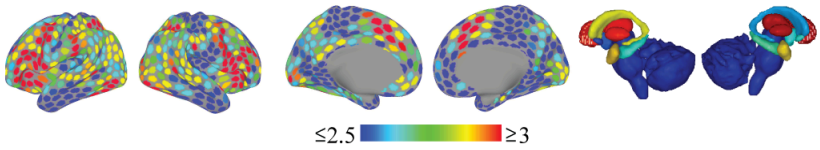

Figure 3.9: Functional diversity scores (Shannon entropy) based on studies in the BrainMap database. Cool colors represent regions of low diversity and warm colors represent regions of high diversity.

3.3.3 Reliability of results

The above results were based on the median of the membership values across bootstrapping iterations. The bootstrapping results give us the ability to define confidence intervals on these estimates. For example, for Figure 3.7, the frequency histograms were generated by considering all 5000 iterations. Each bin shows the median frequency across iterations and an interval around the median. Note that only ROIs with membership values consistently greater than zero were considered in the generation of the histograms.

In addition, for every ROI with nonzero membership value we determined confidence intervals (see Figure 3.10). For instance, consider community OC1 during rest in which approximately 250 ROIs exhibited membership values consistently greater than zero. Figure 3.10 shows the membership values as a function of ROI (reordered in ascending fashion for clarity) and the 95-percentile confidence interval around the median. This figure shows that although there was variability from sample to sample around the median, the estimates reported are reliable.

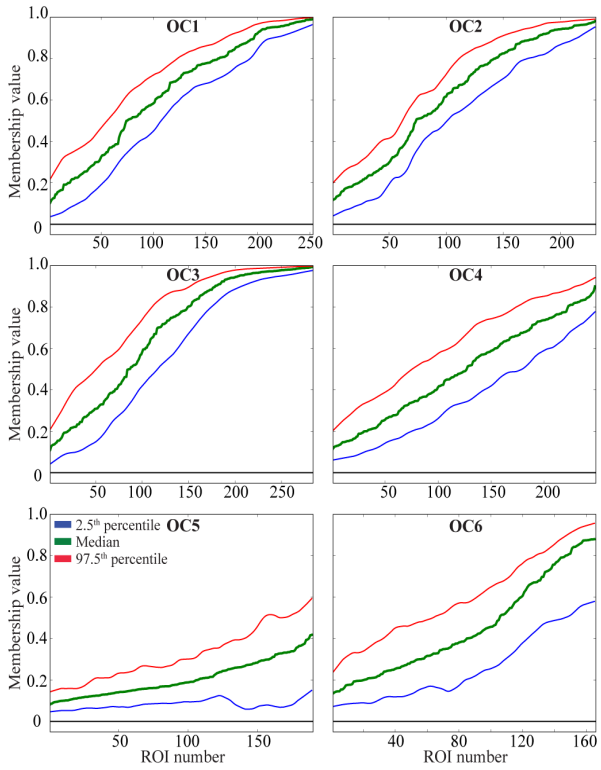

Figure 3.10: Membership values for all ROIs with nonzero values reordered in ascending fashion during rest. Green: median; blue: 2.5th percentile; red: 97.5th percentile.

3.3.4 Does overlapping community structure at the group level reflect that at the participant level?

The results described thus far were based on a resulting group correlation matrix. It is possible, however, that the results were distorted by the implicit averaging associated with our approach. In the extreme, it is possible that each participant had disjoint community structure that differed spatially to some extent, so that the regions with high membership diversity at the group level were artifacts induced by averaging spatially variable subjects. This possibility is increased given recent studies illustrating individual variability in large-scale networks (Gordon et al., 2015;

Mueller et al., 2013). In this section, we outline a series of control analyses aimed at addressing this issue. To anticipate, the results show that dense community overlap is present at the level of individual subjects, and is not an artifact induced by subject averaging.

We ran the mixed-membership algorithm on resting-state data for each of the 94 participants individually. The extracted communities of a few sample participants are shown in Figure 3.11 (for additional results, see Supplementary Figure A.2) to illustrate the quality of overlapping community structure at the individual level. Informal visual comparison with the group results in Figure 3.6 shows good agreement.

To characterize the structure of the estimated communities at the subject level, we matched the labels of the communities between subjects, and examined each communitys distribution of membership values across subjects. The histograms of the estimated communities at the subject level confirm the overlapping structure of the communities rather than exhibiting disjoint structure (Figures 3.12, A.3, A.4, A.5, and A.6). Comparison with the group results (Figure 3.8) also illustrates good agreement in the shape of the distributions (although some differences for OC5 are discernible; but not that this community is also clearly not disjoint).

Another important component of our study was centered on the concept of membership diversity. Therefore, to understand how well group-level membership diversity values reflected membership values at the subject level, we compared the group-level result to the average of subject-level values (Figure 3.13). The strong

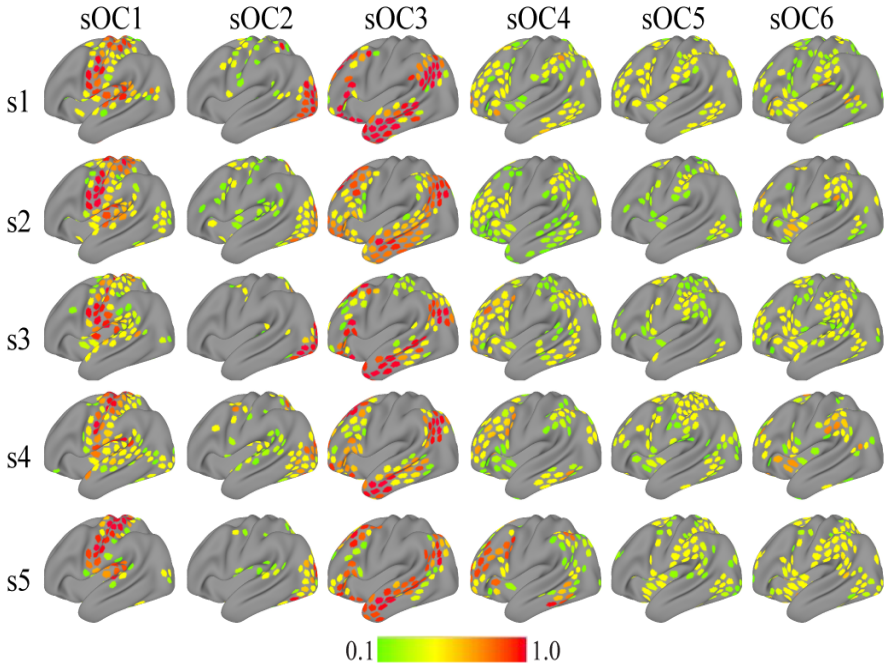

Figure 3.11: Overlapping community organization during rest at the individual level for sample participants. Each row depicts the six overlapping communities extracted with the mixed-membership model per subject (sOC). Membership values are thresholded at 0.1 for illustration.

positive relationship (r=0.817) demonstrates the similarity of the membership diversity reported at the group level to that found for individual participants. To confirm that the high level of correspondence was not driven by only some of the communities, we repeated this comparison for each community, separately (Figure 3.14). The match was excellent for all communities (r$\sim$0.7 or higher) for all communities except community OC5, with a more moderate match (for the complete set of Pearson correlations, see Table 5.1).

Taken together, analyses at the participant level demonstrated that dense overlapping community structure is present at the level of individual participants, and that the group results did not unduly distort patterns observed at the individual

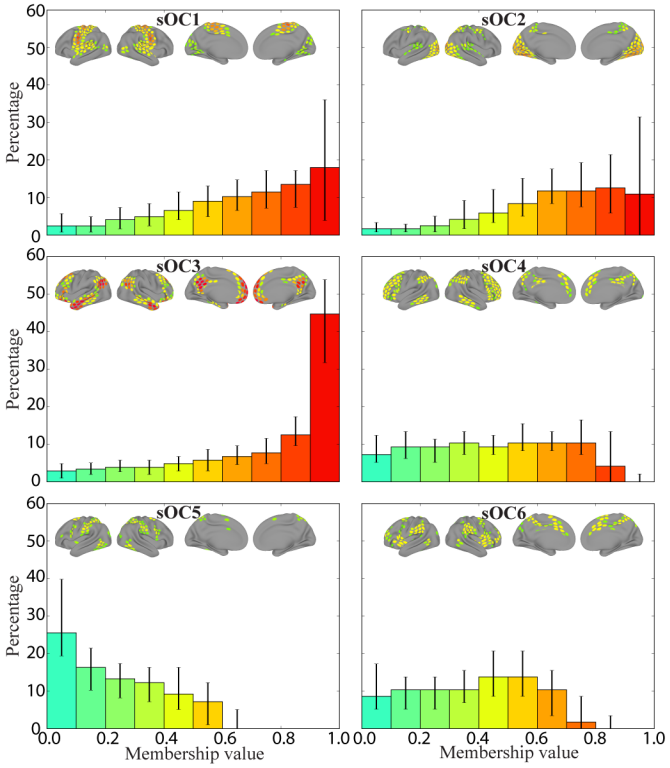

Figure 3.12: Frequency histograms of membership values for overlapping communities estimated at the individual level during rest. Bins show median values across 94 subjects (error bars show the of 25th-75th percentile range). Brain insets show median membership values across subjects (thresholded at 0.1 for illustration). The colors of the bars correspond to the range of membership values shown in the brain insets.

Table 3.1: Correlation between membership diversity at the group and subject level. Only regions with nonzero membership value for communities OC1-OC6 were considered.

Community	OC1	OC2	OC3	OC4	OC5	OC6
Correlation	$r = 0.830$	$r = 0.696$	$r = 0.842$	$r = 0.767$	$r = 0.544$	$r = 0.772$

level.

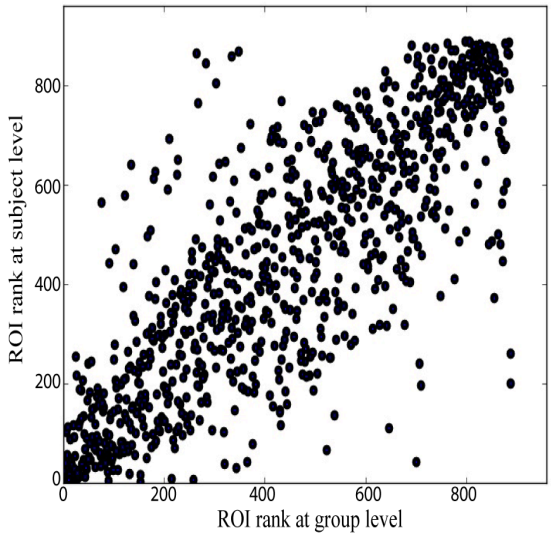

Figure 3.13: Membership diversity values are ranked at the group level and at the subject level, and each region's value is assigned its rank instead of its membership diversity value. The same is done for the average of subject level

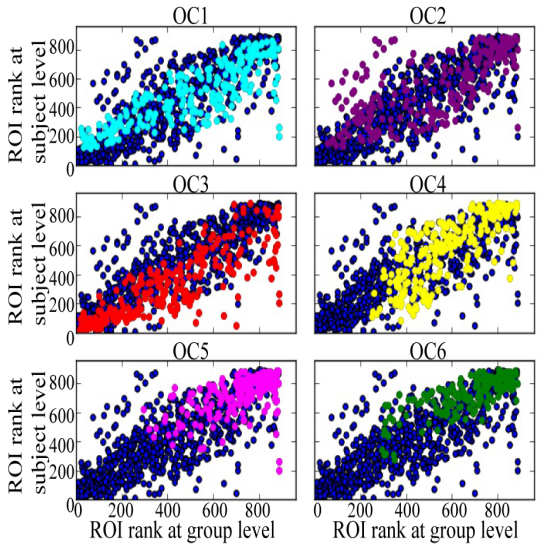

Figure 3.14: Group-level membership diversity of all nodes vs. average of membership diversity at subject level (see also Figure S17). Each different colored node shows a region with robustly nonzero membership value in that community.

3.4 Discussion

We employed a mixed-membership approach to unravel the structure of overlapping large-scale brain networks. Our aims were to compare the community structure of disjoint and overlapping networks and determine the relationship between functional diversity and membership diversity. Below, we discuss the implications of our main findings.

Community structure at rest

Extracting mixed memberships revealed network structure that matched many of the general features of disjoint networks. This was particularly the case for four of the communities (OC1-OC4) that exhibited strong matches (cosine similarity greater than 0.85) with specific disjoint communities (DC1-DC4, respectively). Networks OC5 and OC6 showed somewhat weaker matches to a single disjoint community. In addition to the strongest matching community, several of the mixed communities correlated non-trivially with other disjoint communities; for instance, OC2 and DC5 (0.19), OC3 and DC6 (0.16), and so on. This suggests that mixed communities are not fully captured by a single disjoint community, and that they contain information that is not well described by disjoint communities.

We gained a richer characterization of the networks by considering the distributions of membership values (Figure 3.8; note that the histograms already eliminated ROIs with membership values that were not reliably greater than 0). Data that would be well characterized by a single community would exhibit a membership

distribution concentrated mostly with values close to 1. All mixed communities had non-trivial membership values contributions far from the peak of 1. Inspection of the distributions suggests that all mixed communities would not be well described by a single community. Even for OC3, most of the probability mass (¿60%) was observed in the non-maximal bins. Thus, whereas the mixed-membership approach reproduced many of the general features of the disjoint communities, mixed communities contain information that is not captured by disjoint communities.

Functional diversity

Several groups have suggested that brain regions can be conceptualized via their functional repertoire which is inherently multidimensional (see Anderson et al. (2013); Bzdok et al. (2013); Lancaster et al. (2012); Passingham et al. (2002); Poldrack (2006); Poldrack et al. (2009)). Accordingly, brain regions vary considerably in their functional diversity. In particular, regions such as the anterior insula, as well as parts of lateral and medial frontal cortex, are highly diverse. We conjectured that functional diversity would be linked with membership diversity observed at rest. In other words, regions with high functional diversity should participate in several networks, insofar as they may be engaged with a broad range of regions in the process of participating in diverse functions. Here, we found a positive association between functional and membership diversity, suggesting that the two properties are linked.

Limitations

In this section, we discuss some of the limitations of our study. For one, the choice of number of ROIs was somewhat arbitrary. For instance, Hagmann et al. (2008) employed 998 cortical regions, Power and colleagues ((Power et al., 2012)) used 264 cortical ROIs, and Yeo and colleagues (Yeo et al. (2011)) used 1175 uniformly spaced vertices. In the present study (960 cortical and subcortical ROIs), we chose a relatively large number of regions (and consequently small in size in cortex) because our goal was to investigate overlapping communities; we thus did not want to favor the possibility of overlap by having regions that were large and potentially more functionally heterogeneous. We also note that the cortical ROIs defined here were solely based on the spatial coordinates of the respective surface vertices. Therefore, future work not only should investigate the role of ROI "granularity" (that is, size/number), but also the effect of functional homogeneity in the definition of ROIs. In addition, we did not subdivide subcortical regions into small parcels, which is suboptimal because subcortical regions are heterogeneous and contain subnuclei. Accordingly, future studies will need to re-investigate the overlapping community structure of subcortical regions at finer levels of parcellation.

Another limitation of this study is that it investigated a single overlapping community algorithm. Research on this class of algorithms (Xie et al., 2013) has grown considerably since at least the impactful publication by Palla and colleagues (Palla et al., 2005). Therefore, it will be important to investigate overlapping community structure across a broader range of algorithms; for a different approach than

the one adopted here, see (Yeo et al., 2014). Some of the advantages of the method employed here include its scalability to massive networks given the new methods used for stochastic variational inference (these scale only linearly on the variables of interest, such as the number of node pairs and the number of communities).An important feature of the method is the ability to provide continuous membership values (from 0 to 1) that indicate the extent to which a node belongs to each community, which is in contrast to other methods that treat overlap as binary (present/absent).

Next limitation of the present method is that the results depend on the parameter K, the number of communities. However, the probabilistic nature of the method enables the use of predictive methods to find the best model fit given the data, thereby assuaging this problem somewhat. Nevertheless, a more complete investigation of the overlapping community structure as a function of the number of communities is warranted.

Finally, the employed method does not accept weighted edges, thus requiring a binarization step. Here, we binarized the data with a thresholding method that identifies edges that are consistently among the strongest links for each participant. Nevertheless, in general, methods that avoid binarization may uncover important information in networks (Goulas et al., 2015), and the study of overlapping community structure with weighted networks should be pursued in future studies.

Overall, we conclude that overlapping network methods provide a promising framework to investigate the structure of large-scale brain. While much work has described the brain in terms of disjoint clusters, this type of characterization does not capture the flexible and task dependent mapping between brain regions and

their functions.

Chapter 4: Overlapping communities revealed important differences between community organization during rest and performing task states

4.1 Introduction

In the previous chapter, we showed that characterizing overlapping communities at rest reveals rich structure about the functional networks of the brain, and employing disjoint community detection methods which assume each region belong to single community limits our understanding from brain functional networks.

In this chapter, we sought to characterize the overlapping community structure during task performance, and study how the network structure during rest is potentially altered by task execution. This is important because the large-scale functional structure of the brain during task performance is rarely studied, and less is known about how task performance reconfigures brain's large-scale networks. Critically, it is at times assumed that functional connectivity at rest is affected in minor ways by tasks (Cole et al., 2014). In this view, the activity covariation at rest is only mildly influenced by task execution. While some studies have indeed provided evidence in favor of this idea, an alternative proposal is that tasks alter patterns of functional

connectivity more substantially (Buckner et al., 2013).

At least since the work by Guimera and Amaral (2005*a*,*b*) there has been increased appreciation that particularly well-connected nodes, often called hubs, can be grouped into several distinct sub-types: provincial hubs (well-connected nodes with almost all of their links within a single community), connector hubs (well-connected nodes with at least half of their links within a community), and kinless hubs (hubs with fewer than half of their links within a community). The hub sub-types can be naturally defined by characterizing each node's *bridgeness* (Nepusz et al., 2008), which is defined as the ability of a node to participate with multiple communities simultaneously and "bridge" them together. The different hub sub-types are useful for understanding the general functional organization of networks because each of the defining connectivity patterns lends itself to a "universal role" that does not depend on the type of network being studied (social, technological, or biological). Thus, our next objective was to measure "bridgeness" of each region to extend our understanding of "universal roles" and identify key information processing nodes in the brain. Specifically, finding the bridges in the brain is important because they have the potential to spread signals across multiple communities, thereby performing important roles in distributed processing. Furthermore, we wished to determine how both hub and bridge properties changed during task execution relative to rest.

4.2 Materials and Methods

4.2.1 Datasets

The structural and functional MRI data for this study were obtained from the Human Connectome Project task datasets, with the same set of subjects from Chapter 2 (for details, see section 4.2.1, Glasser et al. 2013; Smith et al. 2013; Van Essen et al. 2013).

Here, we employed HCP's so-called "minimally processed" Emotion and Working mermory functional datasets (Glasser et al., 2013), which included the following preprocessing steps: correction of spatial distortions, motion correction, functional to structural data alignment, bias field correction, and intensity normalization. Cortical data were mapped to the surface (using the Conte69, 32k standard mesh), and subcortical data were analyzed volumetrically. Surface data were spatially smoothed within the surface with a 2 mm kernel. The experiment design of Working memory and Emotion tasks are explained as follows (for details of resting-state dataset see section).

4.2.1.1 Working memory task data

During the working memory task, participants performed an n-back memory task based on a series of centrally present pictures of places, tools, faces and non-mutilated body parts with no nudity. Participants performed two runs of the task, each of which contained four blocks of a 0-back memory task, four blocks of a 2-back

memory task, and four blocks of fixation. On each trial, a stimulus was presented for 2 seconds, followed by a 500-ms interval. Here, we restricted our analysis to the 2-back condition. Each task block lasted 25 seconds; to account for the hemodynamic delay, we employed volumes from 10-25 seconds after block onset. In total, we employed 120 seconds (167 volumes) of data for each participant (2 runs $\times$ 4 blocks/run $\times$ 15 seconds/block). The only processing step applied beyond the HCP pipeline was to regress out (via the 3dDeconvolve program of the AFNI package) 12 motion-related variables: 6 motion parameters estimated from the rigid-body transformation to the reference image acquired at the beginning of the scan and their temporal derivatives (as provided with the HCP data).

4.2.1.2 Emotion task data

The emotion task was based on the original paradigm developed by Hariri and colleagues (Hariri et al., 2000). During each block, participants decided which of two faces (or shapes) on the screen matched the face (or shape) placed above. On each trial, stimuli appeared for two seconds followed by a one second interval. Here, we restricted our analysis to the emotion condition, during which angry or fearful faces were displayed. Each task block lasted 21 seconds; to account for the hemodynamic delay, we employed volumes from 10-21 seconds after block onset. In total, we employed 66 seconds (92 volumes) of data for each participant (2 runs $\times$ 3 blocks/run $\times$ 11 seconds/block). As in the case of the working memory task, the only processing step applied beyond the HCP pipeline was to regress out of the estimated motion

parameters.

4.2.1.3 Data censoring

Because head motion can lead to spurious patterns of co-activation, it is important to preprocess and censor the data so as to minimize potential artifacts. Thus, we calculated the "temporal derivative of variance over voxels" (DVARs; Power et al. (2012)) to gauge the rate of change in the functional MRI signal at each time sample. If DVARs exceeded a threshold of 0.35 at a time point, the data point was removed; for working memory 5.48% of the data were removed, and for emotion 4.32% of the data were removed. This procedure was applied to task data only because the ICA-fixed preprocessed resting-state data have motion parameters "aggressively regressed out" already (Smith et al., 2013). See Supplementary Material B.1.1 for further discussion of the impact of preprocessing steps on resting-state and task data.

Participant censoring

Data from six participants contained time points with DVARs greater than 0.35 in over half of their blocks in working memory or emotion datasets. We therefore dropped these participants (including resting state) from further analysis. Thus, the final analysis used data from 94 participants.

4.2.2 Regions of Interest

The Region of interest (ROIs) set are the same as the 960 ones used in Chapter 2. For details see section 3.2.2. In brief, 941 cortical ROIs were defined by clustering the vertices on the cortical mesh based on their spatial coordinates. 19 subcortical regions of interest were defined anatomically based on Freesurfer's subcortical segmentation (Fischl, 2012).

4.2.3 Functional connectivity

To calculate the Functional connectivity between every pair of regions for each participant and condition [working memory, emotion], we employed the same procedure as what used for resting-state data in Chapter 2 (for details see section 3.2.3). In short, first we used Pearson correlation to find pairwise correlation between each pair of ROIs and binarized each participant's connectivity matrix by setting the top 10% strongest correlations to 1 and all other entries of the correlation matrix to 0. The binarized connectivity matrices were averaged across participants and binarized a second time with the top 10% threshold again.

4.2.4 Detecting disjoint communities

To estimate disjoint communities during performing task, we employed two commonly used disjoint community algorithms: *Infomap* (Rosvall and Bergstrom, 2008) and *modularity maximization* (Clauset et al., 2004). The clustering methods was applied to the averaged adjacency matrix to estimate disjoint communities. Both of

these methods estimate the number of communities as a part of clustering algorithm.

4.2.5 Detecting overlapping communities

To estimate the overlapping communities during task performance, we employed the mixed-membership algorithm which was developed by Gopalan and Blei (Gopalan and Blei, 2013), as explained in section 3.2.5.1.

For completeness, briefly, the algorithm proceeds as follows. Initially, a *membership matrix* is estimated based on the adjacency matrix and the number of communities to extract, K. The membership matrix has size N_{ROIs}-by-K, where N_{ROIs} is the number of ROIs. Each column of the membership matrix defines a community. Each row of the membership matrix is a membership vector, $\boldsymbol{\pi}_i = [\pi_{i,1}, \pi_{i,2}, ..., \pi_{i,K}]$, that indicates the extent to which a region i belongs to each of the K communities. Importantly, the sum of the membership values along a row sum to 1. Formally, element $\pi_{i,j}$ in the membership matrix is the membership value of region$i$ to the j-th community and, for every region i, $\sum_{j=1}^{K} \pi_{ij} = 1$ (Figure 3.2). For details see section 3.2.5.1.

Choosing the number of communities

To determine the number of communities, for each condition, the Bayesian mixed-membership model fit to the data was computed for different numbers of communities, ranging between 2 and 25. To hold the value of k constant across all bootstrapping iterations, we computed the model fit to the data by considering the

entire original dataset ($N = 94$). Figure 4.1 depicts the model fit for rest, working memory and emotion tasks as a function of k. Given that the fit curves were fairly flat between 4-6 communities for both the working memory and the emotion tasks, and having a peak at $k = 6$ for resting state, we fix $k = 6$ across all conditions to facilitate comparison between rest and task conditions. As mentioned in Chapter 2, however, that this does not imply that the value is "correct," as choices of k between 4 and 6 (across datasets) were quite similar; in addition, these values of k are specific to the model employed here, and other models could partition the data optimally with a different number of communities.

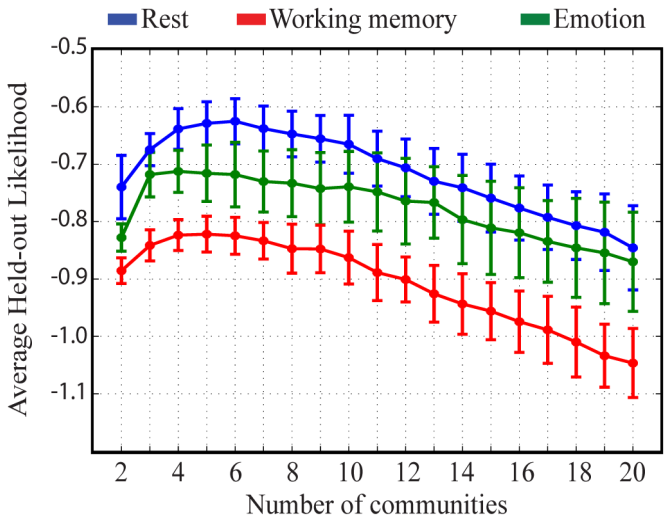

Figure 4.1: Model fit as a function of the number of communities for all conditions (resting-state dataset, working memory, and emotion) investigated. Because the trend continues after $k = 20$, we do not show values $20 < k \leq 25$.

Reproducibility and reliability of the results

To study the reliability of estimated overlapping networks during performing task, we applied bootstrapping (Efron, 1979) to our data by randomly resampling subjects

(with replacement) for a total of 5000 iterations (the same sets of subjects in each iteration as the resting-state data in Chapter 2). Overlapping network analysis was performed independently for each of the 5000 iterations (each sample of 94 participants defined its own functional connectivity matrix). For further details on the method see section 3.2.6.

Membership diversity

We computed the diversity of membership values for each ROI by employing Shannon's entropy measure (Shannon, 2001), in a similar way as section 3.2.8 for working memory and emotion datasets.

4.2.6 Degree

Degree of a region is defined as the number of links attaching it to other regions. Because the mixed-membership utilized a binarized graph, the degree of i-th region was computed as the sum of the values in the i-th row of the binary adjacency matrix minus 1. Formally,

$$D_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} A_{ij} \tag{4.1}$$

where D_i is degree of region i, and A_{ij} is the link value between region i and region j. Having the binary adjacency matrix for group level, if there is a connection between regions i and j, A_{ij}, otherwise it equals 0.

4.2.7 Modularity of overlapping networks

Network modularity is used in the literature to indicate how a network can be subdivided into communities. For example, in the modularity measure defined by Newman for disjoint communities (Newman, 2006), if two regions are connected and belong to the same community, modularity increases, and if two regions are connected but belong to two different communities, modularity decreases. Several modularity methods have been defined for overlapping communities, too (Xie et al., 2013). Here, based on Chen and colleagues (Chen et al., 2010), we extended modularity to the case of the mixed-membership model we employed as follows:

$$Q_c = \frac{1}{2m} \sum_{c} \sum_{i,j \in c} (A_{ij} - \frac{D_i D_j}{2m}) \pi_{i,c} \pi_{j,c} \tag{4.2}$$

where m is the total number of links in the graph, A_{ij} is the edge between node i and j (1 if edge present, 0 otherwise), D_i is the degree of node i (based on the binarized group graph, equation 4.1), and $\pi_{i,c}$ is the membership value of node i in community c. This extension of modularity is straightforward. The part of the equation prior to the product of the membership values, $\pi_{i,c}\pi_{j,c}$, implements the intuitive description of standard (disjoint) modularity provided above; the product of the membership values incorporates the "strengths" into the original formulation (which essentially consider membership as a binary value).

where m is the total number of links in the graph, is the edge between node i and j (1 if edge present, 0 otherwise), is the degree of node i (based on the binarized group graph), and is the membership value of node i in community c. This extension

of modularity is straightforward. The part of the equation prior to the product of the membership values, , implements the intuitive description of standard (disjoint) modularity provided above; the product of the membership values incorporates the "strengths" into the original formulation (which essentially consider membership as a binary value).

4.2.8 Defining Bridges

To identify ROIs acting as bridges, first we normalized both degree and membership diversity to have values between 0 and 1. Bottleneck bridges are regions with high membership diversity and low degree. Hub bridges are regions with high membership diversity and high degree. Thus, bridgeness scores can be defined as:

$$B = \begin{cases} z((1 - degree) \times membership\ diversity) & \text{for bottleneck bridges} \\ z((degree) \times membership\ diversity) & \text{for hub bridges} \end{cases} \tag{4.3}$$

and computed for each bootstrapping iteration, when degree and membership diversity for each iteration are normalized between 0 and 1. $z(.)$ function computes the z score of the argument.

Power and colleagues (Power et al., 2013) argued that studies that use *degree* to identify hubs in resting-state functional networks are problematic because the identified hubs may be due to community size rather than their purported roles in "global" information processing. In other words, if *degree* is the only parameter used to determine if a region is a "hub," the usage is potentially problematic. Note,

however, that here the information processing role of nodes was also based on membership diversity. Thus, when membership diversity is low (say, a region is mostly affiliating with one or two communities), the bridgeness score will be low. In essence, therefore, degree played a role in labeling bridges as "hub bridges" or "bottleneck bridges," which indicates if a bridge between communities is linked to multiple or a few other regions, respectively. In any case, further analyses revealed no evidence of a systematic effect of community size on node taxonomy (see Supplementary Section B.1.2).

4.3 Results

In the next two sections, we report on the overlapping network structure as revealed by task data; specifically working memory and emotion tasks. The subsequent sections then compare properties of overlapping networks of resting-state and task data.

4.3.1 Comparing overlapping communities during rest and task

Although network structure has been extensively studied during taskless states, less is known during task execution. In particular, how network structure is potentially altered by tasks is actively debated. To address this question, we investigated the overlapping network structure of functional MRI data collected during working memory and emotion tasks. Because our goal was to compare network properties during task execution to those found at rest, we did not employ the baseline tasks

conditions; instead we used data for the active conditions alone (working memory: 2-back condition; emotion: matching emotional faces; see Materials and Methods). To aid in the comparison with the results at rest, a value of $K = 6$ was used for both cases.

Overlapping communities are shown in Figure 4.2 for working memory (WM1-6) and Figure 4.3 for emotion (EM1-6). Visually, several of the communities during the two tasks resembled communities found at rest. However, careful inspection of the similarity matrices (Figures 4.4 and 4.5) revealed that tasks altered the observed networks in important ways, too. Consider, for example, the working memory task. From the standpoint of the communities observed at rest, only OC2 displayed a substantial match to a single community during working memory (WM2), as evidenced by the bar plots (see insets). Likewise, for the emotion task, again, only OC1-2 displayed a substantial match to a single community during emotion (EM1-2, respectively).

Examination of the distribution of membership values for individual communities provides further insight into overlapping information. For the working memory task (Figure 4.6), most communities exhibited small, intermediate, and large membership values. Community WM2 was the only one that showed a peak at the bin with largest values, and communities WM4-6 were negatively skewed. For the emotion task (Figure 4.7), no community exhibited a peak near 1, and communities EM4-6 showed negative skew.

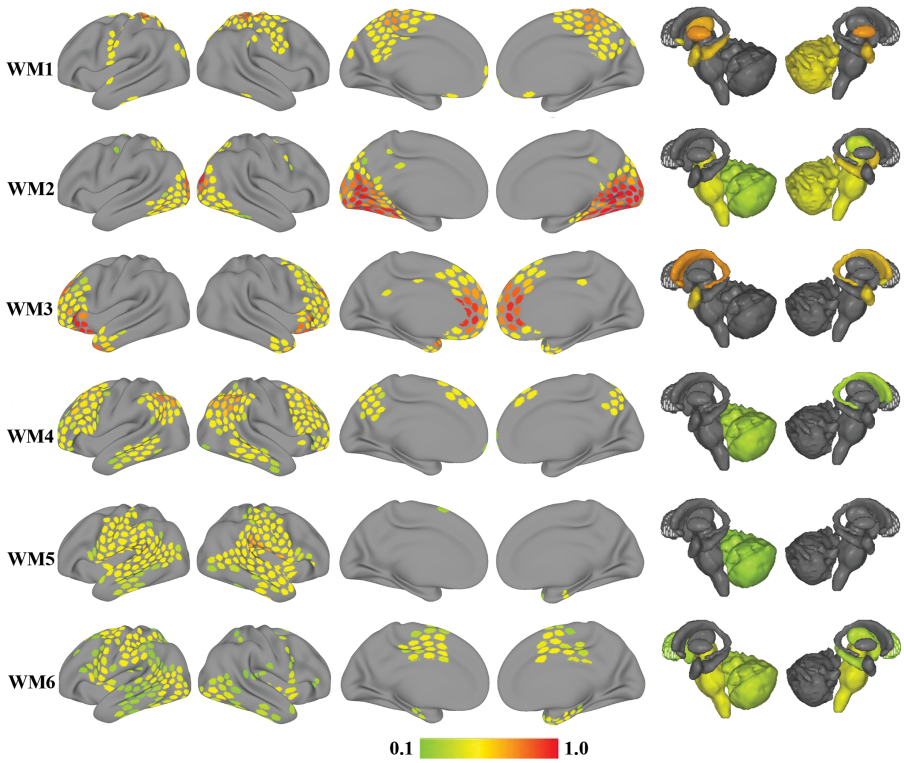

Figure 4.2: Overlapping communities (WM) detected during the working memory task. Each row depicts one of the six overlapping communities extracted with the mixed-membership model. The color of cortical and subcortical regions reflects the median membership value of each region to each community across 5,000 iterations. Membership values are thresholded at 0.1 for illustration.

4.3.2 Disjoint communities during the working memory task

The overlapping community structure for the working memory condition did not reveal clear evidence for a task-negative community. Because this was somewhat surprising, we investigated working memory network organization with two commonly used disjoint community algorithms: *Infomap* (Rosvall and Bergstrom, 2008) and *modularity maximization* (Clauset et al., 2004). Interestingly, neither algorithm extracted a task-negative community during the working memory condition (Figures 4.8 and 4.9). Infomap detected eleven communities (but only six of communities

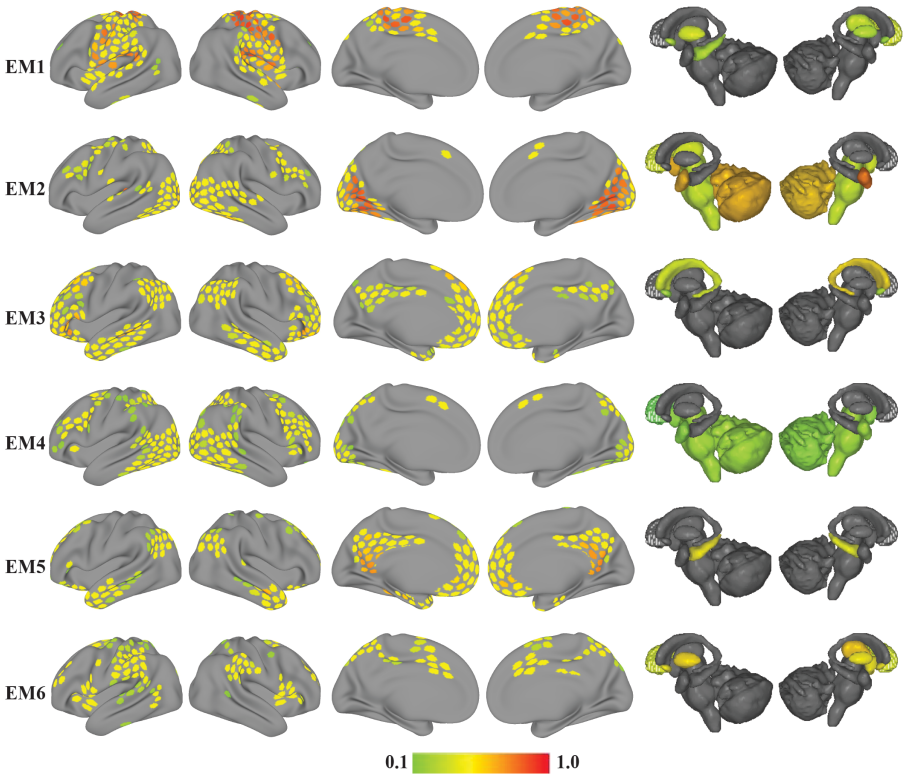

Figure 4.3: Overlapping Communities (EM) detected during the emotion task. Each row depicts one of the six overlapping communities extracted with the mixed-membership model. The color of cortical and subcortical regions reflects the median membership value of each region to each community across 5,000 iterations. Membership values are thresholded at 0.1 for illustration.

had more than ten regions; none of the regions in the five "small communities" was part of the standard task-negative community) and modularity maximization detected six communities (but one community had only one region).

4.3.3 Modularity of overlapping networks

Further insight into the changes in network structure linked to tasks states can be gained by studying modularity. Conceptualizing networks as inherently overlapping structures highlights their non-modular structure. Nevertheless, modularity is not all-or-none, so quantifying it provides a measure of the extent to which signals

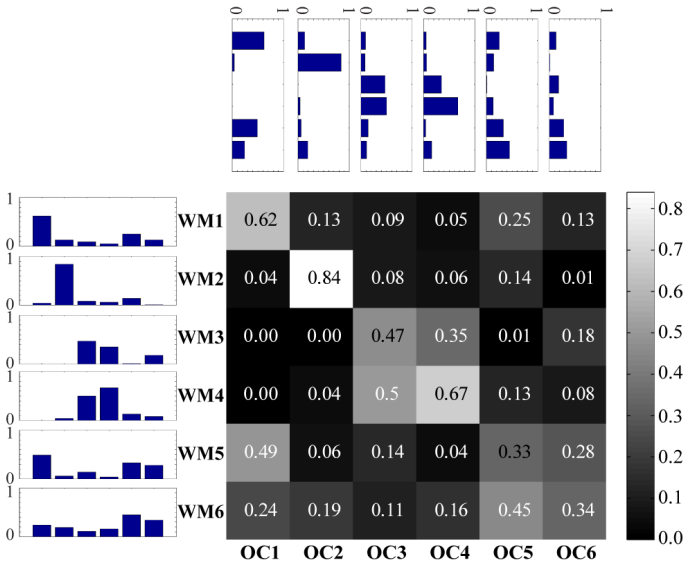

Figure 4.4: Cosine similarity (equation 3.1) between overlapping communities at rest (OC1-OC6) and during the working memory task (WM1-WM6). The matrix displays the median cosine similarity between each WM and OC network across 5,000 iterations. Given that the community membership vectors do not contain negative values, the cosine similarity scores range from zero (orthogonal) to one (identical). Side and top insets represent similarity scores as bar plots across rows and columns, respectively.

potentially can flow between communities. We defined a measure of overlapping community modularity based on membership values (Materials and Methods), and a modularity score was computed for each community during the rest, working memory, and emotion conditions. Modularity (Figure 4.10) was clearly highest during resting state (mean and standard deviation: 0.3180.016), and decreased for both tasks (working memory: 0.1830.018; emotion: 0.2110.015) for which modularity scores were fairly similar. Importantly, all modularity scores of individual communities during working memory and emotion were lower than values observed at rest, showing that the reduction was not driven by changes to one or a just a few communities. Additional analyses in Supplementary Material (Section Supplemen-

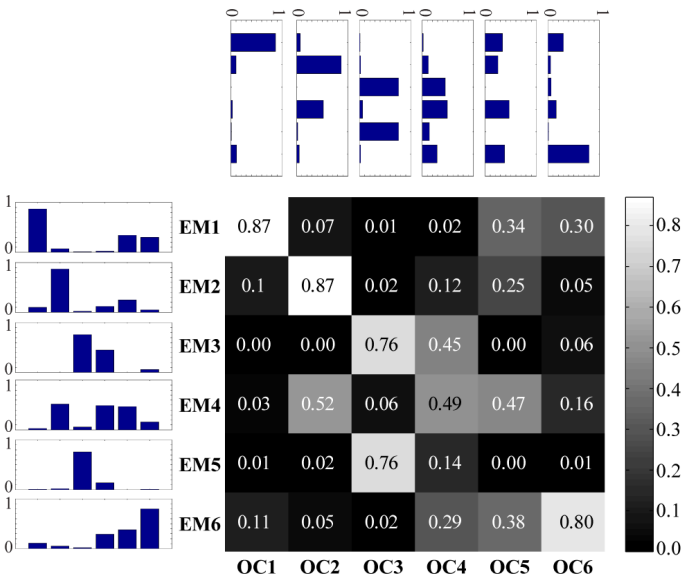

Figure 4.5: Cosine similarity (equation 3.1) between overlapping communities at rest (OC1-OC6) and during the emotion task (EM1-EM6). The matrix displays the median cosine similarity between each EM and OC network across 5,000 iterations. Given that the community membership vectors do not contain negative values, the cosine similarity scores range from zero (orthogonal) to one (identical). Side and top insets represent similarity scores as bar plots across rows and columns, respectively.

tal B.1.3) show that these findings also hold when the number of communities, K, varies across conditions.

4.3.4 Node taxonomy: hubs and bridges during rest and task conditions

An additional goal of this study was to understand how regions potentially participate in "signal communication" so as to better characterize the overlapping network structure of brain networks. More broadly, characterizing node properties within the overall network topology can help elucidate the information processing roles played by nodes (Guimera and Amaral, 2005a,b). The mixed-membership approach em-

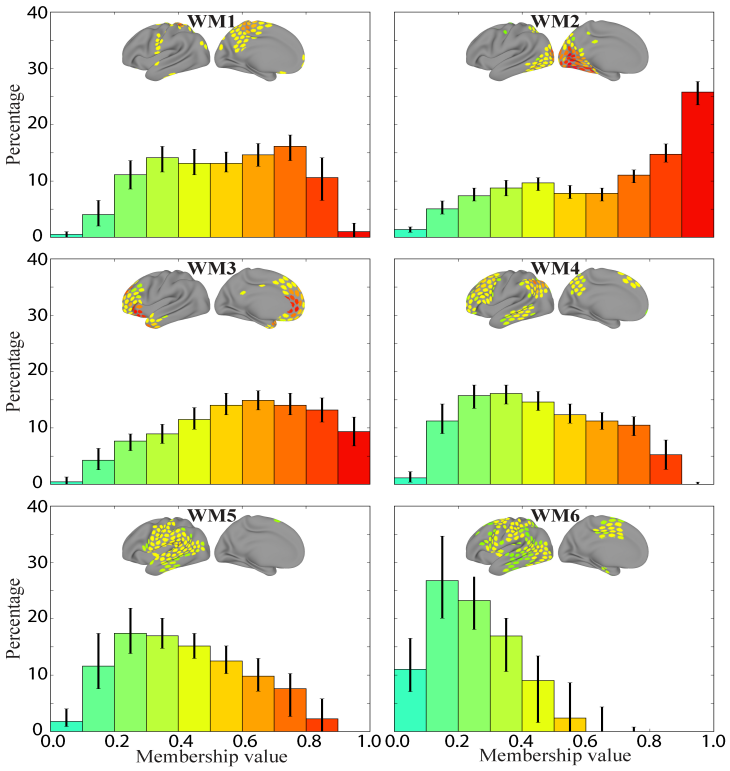

Figure 4.6: Frequency histograms of membership values for each of the six overlapping communities during the working memory task. Each histogram depicts the median value in each bin for that community across 5,000 iterations (error bars show the of 25th-75th percentile range). The colors of the bars correspond to the range of membership values shown in the brain insets.

ployed here is useful because each node is characterized with a set of probability-like values that characterize its participation across all networks simultaneously.

To characterize a node's functional role, we employed the *membership diversity* measure for overlapping communities described above, which captures the extent to which a node participates in multiple communities. Our reasoning was that nodes with high *membership diversity* may function as important "bridges" by facilitating communication across multiple communities. As stated previously, a node's *membership diversity* was quantified by the Shannon entropy of its membership values (section 3.2.8). We also considered the *degree* of a node, a standard graph measure

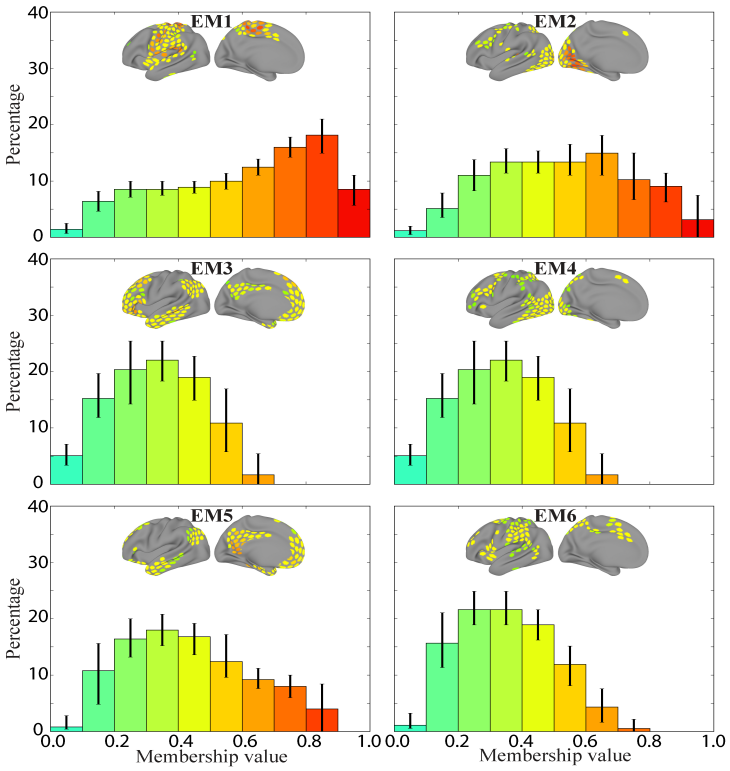

Figure 4.7: Frequency histograms of membership values for each of the six overlapping communities during the emotion task. Each histogram depicts the median value in each bin for that community across 5,000 iterations (error bars show the of 25th-75th percentile range). The colors of the bars correspond to the range of membership values shown in the brain insets.

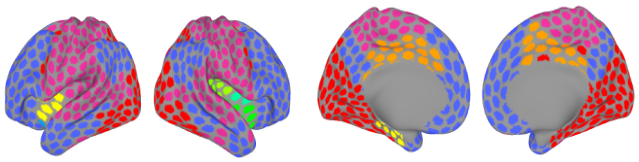

Figure 4.8: Disjoint communities during the working memory task estimated by the Infomap algorithm.

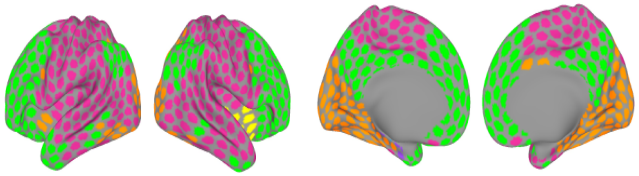

Figure 4.9: Disjoint communities during the working memory task estimated by modularity maximization.

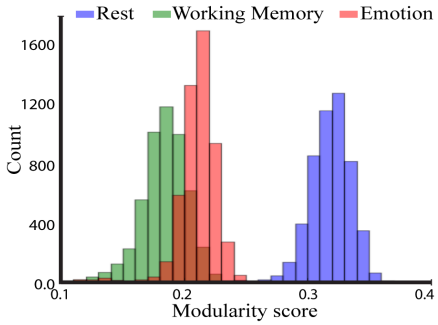

Figure 4.10: Modularity scores of overlapping communities for rest and both tasks. The histograms depict the whole-brain modularity scores across 5,000 iterations (each modularity score was obtained by summing modularity scores across communities).

that indicates how highly connected the node is to all other nodes (Materials and Methods).

degree and *membership diversity* capture different aspects of node function. For example, a region with high *degree* is connected to many regions and a region with low *degree* is connected to a small number of regions. *membership diversity* indexes a different aspect of network structure. For instance, a region with a high membership value for one community (it participates highly in that community) and low membership values elsewhere would have low *membership diversity* because it participates mostly within a single community; in contrast, a region with intermediate membership values across multiple communities would have high *membership diversity* because it participates in multiple communities. Thus, *degree* helps measure the extent to which a region is a "hub" (Guimera and Amaral (2005*a*,*b*); see Section 4.4.2 for further discussion), and *membership diversity* indicates the extent to which a region is a cross-community "bridge" (Nepusz et al., 2008; Yu et al., 2007). Combining these two measures leads to four general classes of regions:

- *Locally connected regions* (low degree/low membership diversity) are not highly connected and communicate primarily within a single community;
- *Local hubs* (high degree/low membership diversity) are highly connected regions that communicate primarily within a community;
- *Bottleneck bridges* (low degree/high membership diversity) are regions with few connections that span multiple communities;
- *Hub bridges* (high degree/high membership diversity) are highly connected regions with connections that span multiple communities.

We were particularly interested in investigating the distribution of the last two types of node above for both rest and task datasets. For visualization, we sorted the regions based on their bridgeness score in each condition (Figures 4.11 and 4.12; Materials and Methods). Notably, several properties observed at rest were altered during task execution. See Discussion for further elaboration and Supplementary Section B.1.2 for additional analyses.

4.3.5 Reliability of results

The above results were based on the median of the membership values across bootstrapping iterations. The bootstrapping results give us the ability to define confidence intervals on these estimates. For example, for Figures 4.6 and 4.7, the frequency histograms were generated by considering all 5000 iterations. Each bin shows the median frequency across iterations and an interval around the median. Note that only ROIs with membership values consistently greater than zero were considered in the generation of the histograms.

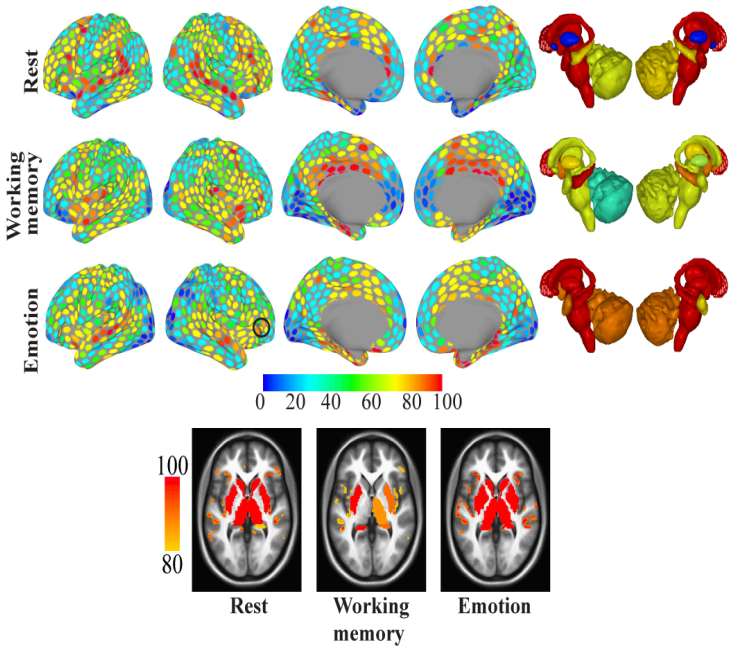

Figure 4.11: Bottleneck bridges. Bridgeness scores for each region and condition (top: resting state, middle: working memory, bottom: emotion). Colors indicate the percentile of the ROI's median score across 5000 iterations (for example: regions colored red had bridgeness scores around the 90th percentile or above). Black contours indicate regions discussed in the text. Bottom row: Horizontal slices at illustrating strong bridges in the anterior insula.

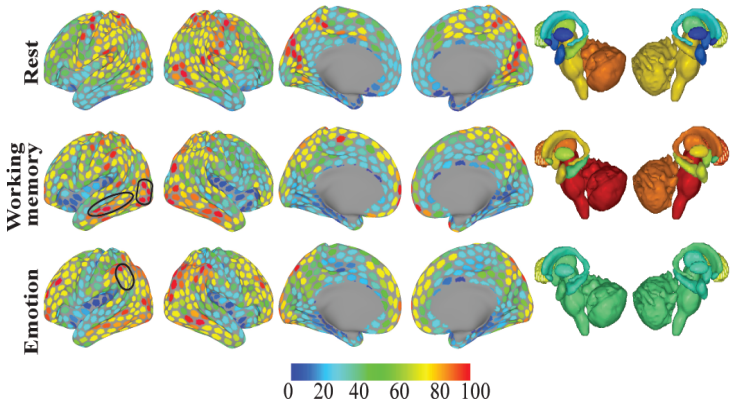

Figure 4.12: Hub bridges. Bridgeness scores for each region and condition (top: resting state, middle: working memory, bottom: emotion). Colors indicate the percentile of the ROI's median score across 5000 iterations (for example: regions colored red had bridgeness scores around the 90th percentile or above). Black contours indicate regions discussed in the text.

In addition, for every ROI with nonzero membership value, and across all three conditions, we determined confidence intervals (see Figures 4.13 and 4.14). Figures 4.13 and 4.14 show the membership values as a function of ROI (reordered in ascending fashion for clarity) and the 95-percentile confidence interval around the median. These figures show that although there was variability from sample to sample around the median, the estimates reported are reliable.

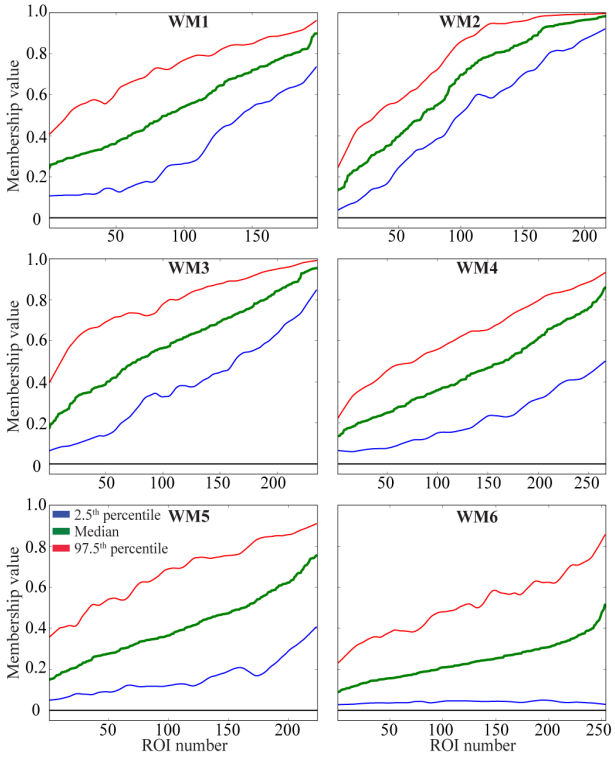

Figure 4.13: Membership values for all ROIs with nonzero values reordered in ascending fashion during the working memory task. Green: median; blue: 2.5th percentile; red: 97.5th percentile.

4.4 Discussion

In this chapter we employed mixed membership community detection method to characterize the overlapping community structure during performing tasks and in-

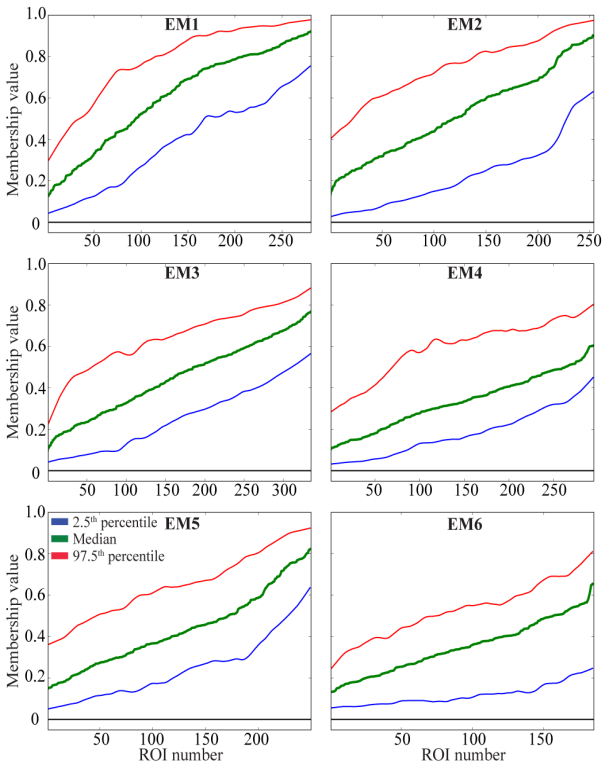

Figure 4.14: Membership values for all ROIs with nonzero values reordered in ascending fashion during the emotion task. Green: median; blue: 2.5th percentile; red: 97.5th percentile.

vestigate how performing task reconfigures the communities at rest condition. Also, we studied the distribution of "bridges" in the brain, including bottleneck and hub bridges, during both rest and tasks. Below, we discuss the implications of our main findings.

4.4.1 How do tasks alter the functional connectivity landscape?

Buckner and colleagues (Buckner et al., 2013) asked the following question: Do networks studied during the resting state capture fundamental units of organization or should "rest" be considered just another arbitrary task state? Both sides of this debate are represented in the literature (for discussion and references, see Cole et al.

(2014)). To investigate this question, we compared mixed networks observed during rest, as well as working memory and emotion tasks.

Although similarities were apparent between rest and task communities, important changes were observed. For example during the working memory task, only two communities that were associated with sensory and sensorimotor aspects (Figure 4.2: WM1 and WM2) were fairly well represented by a single community observed at rest (Figure 3.6: OC1 and OC2, respectively; this was particularly the case for WM2/OC2). The remaining communities correlated nontrivially with two or more communities observed at rest. Likewise, during the emotion task, several communities overlapped with multiple communities observed at rest. Our findings support the idea that considerable reorganization is observed during specific tasks, and that it may be prudent to consider "rest" as a particular task state.

The relationship between networks observed during the resting state and task states was also investigated by determining the modularity structure of the communities. Quantifying modularity provides a measure of the extent to which signals potentially can flow between communities. We can thus consider the inverse of modularity as an index of communicability. Communicability was lowest during resting state and increased for both tasks (Figure 4.10). Importantly, increases in communicability were observed across multiple communities, and were not limited to specific cases, such as decrease only in visual or sensorimotor communities. It thus appears that, during the two tasks studied here, coordinated activity between regions that is important for task execution shapes the observed networks by increasing inter-region integration hence, decreasing modularity.

4.4.2 Characterizing different types of bridges

The framework of overlapping communities offers a natural way to discover nodes that participate across multiple communities - "bridges." Here, we combined node *membership diversity* (which captures the extent to which a node participates in multiple communities) with node *degree* to investigate two types of bridges: *bottleneck bridges*, which are regions with relatively few connections that span multiple communities; and *hub bridges*, which are regions with relatively many connections that span multiple communities. Our approach was to determine bridgeness scores for all ROIs and not adopt an arbitrary threshold, so that the spatial distribution of bridgeness could be better appreciated.

At rest, several bottleneck bridges were found in the prefrontal cortex. These included regions in dorsolateral and more inferior prefrontal cortex. At rest, *hub bridges* were not prominently found in the prefrontal cortex. Notable changes were observed during task execution, some of which we comment on here. During the working memory task, several regions in occipital cortex showed high hub bridge scores. Whereas the same regions were also well connected at rest (they behaved as *localhubs*, which are highly connected regions that communicate primarily within a community), they diversified their participation across communities during the working memory task (and to some extent during the emotion task), thus increasing *bridgeness*. This is interesting in light of the fact that the task required participants to hold in mind information about multiple types of visual stimuli (places, tools, faces, and body parts), and suggests that working memory performance is

characterized by the participation of visual cortex in multiple large-scale networks (see Sreenivasan et al. (2014)). This is also evidenced by the stronger *hub bridges* observed in ventral temporal cortex (which were not prominent during rest).

During the emotion task, *hub bridges* in parietal cortex were stronger more inferiorly in the vicinity of the angular gyrus, whereas they were stronger more medially/superiorly during rest in the superior parietal lobule and the vicinity of the intraparietal sulcus. Furthermore, strong *bottleneck bridges* were not prominent in dorsal prefrontal regions, and were found instead more inferiorly (especially on the right hemisphere). These findings resonate with the roles attributed to the angular gyrus (Seghier, 2013) and inferior frontal regions. Interestingly, sites in the inferior frontal gyrus have been consistently reported in many emotion tasks and implicated in "emotional salience," and a recent meta-analysis has identified the inferior frontal gyrus (on the right hemisphere) as a "hub" for emotional processing (Kirby and Robinson, 2015). Our findings suggest that this region may function as a bottleneck bridge, that is, a region that is not necessarily highly connected but one that participates across several communities. In addition, the anterior insula behaved as a strong *bottleneckbridge*, a role that was also observed at rest (but not during working memory).

Our analysis revealed that multiple subcortical areas play notable roles as bridges. Notably, the amygdala, caudate, putamen, thalamus, and hippocampus were found to be strong *bottleneck bridges* at rest. The cerebellum behaved as a strong hub bridge at rest (especially on the left), showing that it not only had high membership diversity but also high connectivity (*degree*) overall. During the

emotion task, all subcortical regions behaved as strong *bottleneck bridges*. During the working memory task, the brainstem and the cerebellum, as well as the right caudate and right thalamus, behaved as strong *hub bridges*, while the putamen and hippocampus behaved as strong *bottleneck bridges*. Combined, these findings suggest that subcortical regions play important roles in the flow of information at both rest and during specific tasks (but see next section).

Several previous studies of large-scale functional networks did not account for potential contributions of subcortical areas, an omission that has contributed to a cortico-centric view of networks. But it is well documented in the literature that many subcortical areas have massive connectivity with cortex and play part in important cortico-subcortical circuits, including areas such as the striatum, thalamus, amygdala, and cerebellum (Alexander et al., 1986; Amaral et al., n.d.; Jones, 2006; Middleton and Strick, 2000; Sherman and Guillery, 2013).

Finally, we note that nodes typically associated with the task-negative network did not have high *hub bridgeness* scores. This is in contrast to reports based on *degree* that suggest that they are "globally" connected regions (for example, Cole et al. (2010); Tomasi and Volkow (2011)). Here, *degree* played a role in the labeling of bridges as "hub bridges" or "*bottleneck bridges*," which only indicates if a bridge between communities is linked to multiple or a few other regions, respectively.

In conclusion, we illustrated that task performance substantially altered the structure of functional connectivity across brain regions, and enhanced communicability across the brain (that is, modularity decreased during tasks relative to rest). We were also able to study the distribution of "bridge" nodes, including bottleneck

and hub bridges. Task performance altered the role of regions in important ways. We conclude that overlapping network methods provide a promising framework to investigate the structure of large-scale brain networks during both rest and tasks states.

Chapter 5: Intersubject brain network organization during dynamic anxious anticipation

5.1 Introduction

Imagine yourself reclining on a dentist's chair. Most of us experience an aversive reaction to the onset of the drill, and wait anxiously for the moment the dentist will finish checking it and move it toward our mouth. With mouths open, we wait for the drill to make contact with our teeth and brace ourselves for the aversive sensation that may ensue upon contact. This scenario illustrates "anxious anticipation" periods during which negative events will be possibly experienced.

Understanding the neural basis of anxious anticipation is important not only from a basic research perspective, but because aberrant responses to uncertain future negative events are believed to be central to anxiety disorders (Fox and Kalin, 2014; Grupe and Nitschke, 2013). A growing literature of both non-human and human research indicates that anticipatory processing of negative events engages multiple brain regions (Davis et al., 2010; Grupe and Nitschke, 2013; Tovote et al., 2015), including medial prefrontal cortex, insula, and orbitofrontal cortex, cortically. Subcortically, implicated regions include the amygdala, periaqueductal gray

(PAG), and the bed nucleus of the stria terminalis (BNST); the latter has received considerable attention in the past decade (see (Davis et al., 2010; Fox et al., 2015).

Despite recent progress, important questions remain largely unanswered. First, how does anxious anticipation engage and reorganize large-scale brain networks? We propose that emotional processing needs to be characterized at the level of distributed networks (Pessoa, 2017), and not just at the level of evoked responses in specific brain regions (such as the amygdala or BNST). Along these lines, Hermans and colleagues (Hermans et al., 2011) described greater salience-network connectivity during periods of anxiety associated with watching an aversive movie. In a previous study, McMenamin et al. (2014) investigated network interactions when participants were in either threat (unpredictable mild shocks might be administered) or safe (no shocks possible) blocks (McMenamin et al., 2014). It was found that the salience network exhibited a transient increase (following block onset) in network cohesion (that is, within-network functional connections increased) followed by decreased cohesion during a subsequent sustained period. This study thus revealed changes to network organization during transient and sustained periods of threat.

Second, anxious anticipation is inherently temporal. Although previous studies have investigated how brain responses are sensitive to threat proximity (Grupe et al., 2012; Mobbs et al., 2010; Somerville et al., 2010), almost nothing is known about how patterns of brain co-activation (thus networks) change during dynamic manipulations of threat.

To address these important gaps in the literature, we modulated threat dynamically during functional MRI scanning. Two circles moved on the screen for

periods of 60 seconds, sometimes moving closer and sometimes moving apart (Figure 5.1). If they touched, an unpleasant shock was delivered to the participant. We sought to determine how network properties changed during periods of approach (circles moving closer) and periods of retreat (circles moving apart). As in McMenamin et al. (2014), we studied a set of regions spanning the salience, executive, and task-negative networks, given their involvement in cognitive and emotional processing (Yeo et al., 2011). In addition, we investigated a number of subcortical regions, many of which feature prominently in studies of the emotional brain (Pessoa, 2017).

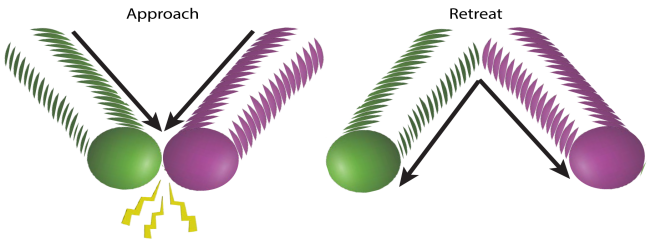

Figure 5.1: **Stimuli paradigm.** To create anxious states, over a period of 60 seconds, two circles with different colors moved around on the screen with some degree of randomness. When they collided with each other, an unpleasant mild electric shock was delivered.

We investigated the questions of interest within the framework of intersubject correlation analysis (Hasson et al., 2004). In this framework, time series data from voxels or regions of interest (ROI) are correlated across participants to determine "interpersonal synchronization" (Figure 5.2A). Intersubject correlation is believed to reflect synchronization of mental states that are not simply explained by common evoked responses to perceptual features or cognitive demands (Ames et al., 2015; Lahnakoski et al., 2014; Nummenmaa et al., 2012, 2014).

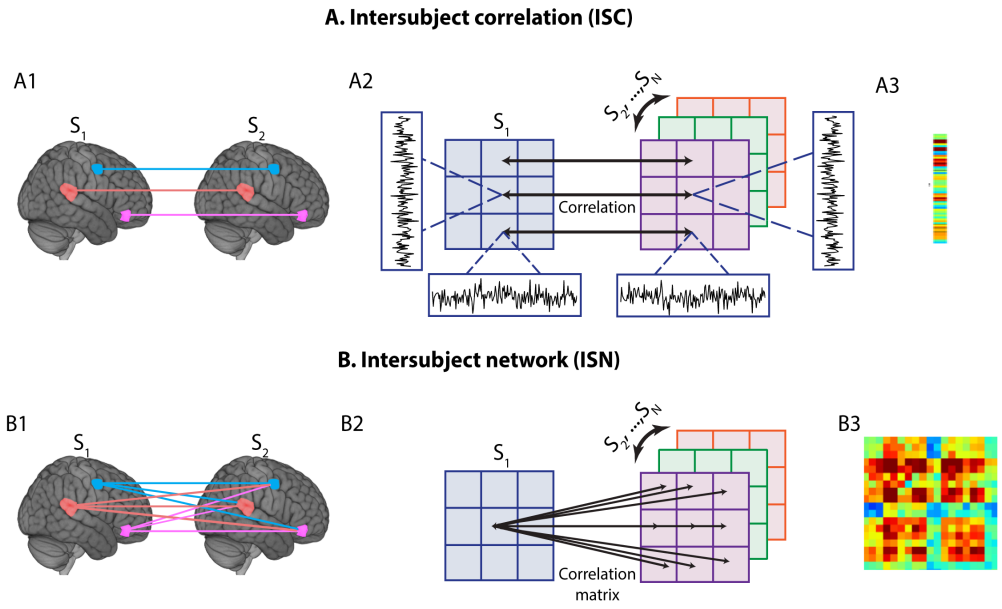

Figure 5.2: **Intersubject correlation (ISC) and network analysis (ISN). A1)** In ISC the correlation between the same region across different participants' brains is calculated. **A2)** To calculate ISC, for each voxel or region of interest (ROI), the time series "left out" subject (S_1) extracted, and correlated with the average time series of all other subjects (S_2, S_N), for the same voxel/ROI. **A3)** To calculate the group level ISC, the results from A2 across all subjects are averaged. This creates a vector that contains the correlation of every voxel with itself (across participants). **B1)** In ISN the correlation between all pairs of ROIs across different brains is calculated. **B2)** To calculate ISN, for each voxel or ROI, the time series of a "left out" subjected (S_1) is extracted and its correlation with the average time series across other subjects ($S_2, , S_N$) is calculated. **B3)** The group level ISN is a $N_{ROIs} \times N_{ROIs}$ matrix, which shows the average of ISN from all subjects. Note that the vector in A3 corresponds to the diagonal of the matrix in B3, illustrating that intersubject networks provide a richer characterization of time series relationships.

Overall, our approach allowed us to test several questions about the brain basis of anxious anticipation. 1) How does dynamic threat reorganize the functional organization of large-scale brain networks? 2) How do network properties evolve during periods of threat approach and retreat? In particular, we sought to test the hypothesis that network organization evolves temporally during threat processing and that, for instance, salience-network cohesion increases/decreases with

threat approach/retreat (Pessoa and McMenamin, 2016). 3) What is the relationship between cortical and subcortical regions important for threat processing during dynamic threat? 4) How are the amygdala and BNST involved in anxious anticipation? This last question is important because, in particular, it is unclear if/how the amygdala is involved in conditions involving relatively prolonged and uncertain threat periods. The role of the BNST also remains unclear, as some have advocated that threat unpredictability is a critical determinant of its involvement (Alvarez et al., 2011).

5.2 Materials and Methods

5.2.1 Data and Experiment Design

Participants

Eighty-five participants (41 females, ages 18-40 years; average: 22.62, STD: 4.85) with normal or corrected-to-normal vision and no reported neurological or psychiatric disease were recruited from the University of Maryland community (of the original sample of 93, data from 7 subjects were discarded due to technical issues during data transfer [specifically, field maps were lost] and 1 other subject was removed because of poor structural-functional alignment). The project was approved by the University of Maryland College Park Institutional Review Board and all participants provided written informed consent before participation.

Procedure and Stimuli

To create anxious states, two circles with different colors moved around on the screen randomly. When they collided with each other, an unpleasant mild electric shock was delivered. Overall, the proximity and relative velocity of the circles were used to influence the threat state. The position of each circle (on the plane), $\mathbf{x}_t$, was defined based on its previous position, $\mathbf{x}_t$, plus a random displacement, $\Delta\mathbf{x}_t$:

$$\mathbf{x}_t = \mathbf{x}_t + \Delta\mathbf{x}_t \tag{5.1}$$

The magnitude and direction of the displacement was calculated by combining a normal random distribution with a momentum term to ensure motion smoothness, while at the same time remaining (relatively) unpredictable to the participants. Specifically, the displacement was updated every 50 ms as follows:

$$\Delta\mathbf{x}_t = (1 - c)\Delta\mathbf{x}_t + cN(0, \sigma) \tag{5.2}$$

where $c = 0.2$ and $N(0, \sigma)$ indicates the normal distribution with 0 mean and standard deviation 1. The position and amount of displacement of each circle was updated independently.

Visual stimuli were presented using PsychoPy (http://www.psychopy.org/) and viewed on a projection screen via a mirror mounted to the scanner's head coil. The total experiment included 6 runs, each of which had 6 blocks (3/85 participants had only 5 runs). In each block, the circles appeared on the screen and moved around for 60 seconds; blocks were preceded by a 15-second blank screen. Each run ended with 7 seconds of a blank screen.

To ensure that the effects of threat proximity and approach were uncorrelated, half of the blocks in each run were the temporally reversed versions of the other blocks in that run. Temporally reversing the stimulus trajectories guarantees that that proximity and approach are uncorrelated because reversing time changes the sign of the approach effect (that is, approach becomes retreat).

In each of the 6 runs the circles collided a total of 8 times in 4 out of the 6 blocks (3 shocks maximum per block); each collision resulted in the delivery of an electric shock. The 500-ms electric shock was delivered by an electric stimulator (Coulbourn Instruments, PA, USA) to the fourth and fifth fingers of the non-dominant left hand via MRI-compatible electrodes. To calibrate the intensity of the shock, each participant was asked to choose his/her own stimulation level immediately prior to functional imaging, such that the stimulus would be "highly unpleasant but not painful." After each run, participants were asked about the unpleasantness of the stimulus in order to re-calibrate shock strength, if needed.

To optimize the experimental design, 10,000 candidate stimuli trajectories and block orders were generated. We then selected 6 runs which minimized collinearity between all predictors of interest (see below), measured as the sum of respective variance inflation factors (Neter et al., 1996).

Stimulus Conditions

We defined two conditions, "approach" and "retreat," based on whether the circles were moving toward or away from each other. Time points were only considered

for analysis if the Euclidian distance between the circles was at most 75% of the maximum distance that the circles would exhibit during the whole experiment; otherwise, the data were not employed in the analysis. The rationale behind this was that, when the circles were far from each other, participants reported that they did not really pay as much attention to them. Therefore, we reasoned that the analysis should focus on the time points during which the circles were in (relative) closer proximity to each other. Investigation of the exploratory set revealed that the particular cutoff was not critical for the effects investigated (in the exploratory set only) and that values at least between 65% and 85% were adequate; based on the exploratory set results we chose a cutoff value of 75%.

MRI data acquisition

Functional and structural MRI data were acquired using a 3T Siemens TRIO scanner with a 32-channel head coil. First, a high-resolution T2-weighted anatomical scan using Siemens's SPACE sequence (0.8 mm isotropic) was collected. Subsequently, we collected 457 functional EPI volumes using a multiband scanning sequence (Feinberg et al., 2010) with TR = 1.0 sec, TE = 39 ms, FOV = 210 mm, and multiband factor = 6. Each volume contained 66 non-overlapping oblique slices oriented 30 clockwise relative to the AC-PC axis (2.2 mm isotropic). In addition, a high-resolution T1-weighted MPRAGE anatomical scan (0.8 mm isotropic) was collected. Finally, double-echo field maps (TE1 = 4.92 ms, TE2 = 7.38 ms) were acquired with acquisition parameters matched to the functional data.

5.2.2 Functional MRI preprocessing

To preprocess the functional and anatomical MRI data, a combination of packages and in-house scripts were used. The first three volumes of each functional run were discarded to account for equilibration effects. Slice-timing correction (with AFNI's 3dTshift) used Fourier interpolation to align the onset times of every slice in a volume to the first acquisition slice, and then a six-parameter rigid body transformation (with AFNI's 3dvolreg) corrected head motion within and between runs by spatially registering each volume to the first volume.

Skull stripping determines which voxels are to be considered part of the brain and, although conceptually simple, plays a very important role in successful subsequent co-registration and normalization steps. Currently, available packages perform sub-optimally in specific cases, and mistakes in the brain-to-skull segmentation can be easily identified. Accordingly, to skull strip the T1 high-resolution anatomical image (which was rotated to match the oblique plane of the functional data with AFNI's 3dWarp), we employed six different packages (ANTs (Avants et al., 2011), AFNI (Cox, 1996): http://afni.nimh.nih.gov/, ROBEX (Iglesias et al., 2011): https://www.nitrc.org/projects/robex), FSL: http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/, SPM: http://www.fil.ion.ucl.ac.uk/spm/, and Brainsuite (Shattuck and Leahy, 2002) and employed a "voting scheme" as follows: based on T1 data, a voxel was considered to be part of the brain if 4/6 packages estimated it to be a brain voxel; otherwise the voxel was not considered to be brain tissue (for 6 subjects whose T1 data were lost due to issues during data transfer, the T2 image was used instead and only the

ANTs package was used for skull-stripping).

Subsequently, FSL was used to process field map images and create a phase-distortion map for each participant (bet and $fsl\_prepare\_fieldmap$). FSL's epi_reg was then used to apply boundary-based co-registration to align the unwarped mean volume registered EPI images with the skull-stripped anatomical image (T1 or T2) along with simultaneous EPI distortion-correction (Greve and Fischl, 2009).

Next, ANTS was used to determine a nonlinear transformation that mapped the skull-stripped anatomical image (T1 or T2) to the MNI152 template (interpolated to 1-mm isotropic voxels). Finally, ANTS combined the nonlinear transformations from co-registration/unwarping (from mapping mean functional EPI images to the anatomical T1 or T2) and normalization (from mapping T1 or T2 to the MNI template) into a single transformation that was applied to map registered functional volumes of functional data to standard space (interpolated to 2-mm isotropic voxels). In this process, ANTS also utilized the field maps to simultaneously minimize EPI distortion.

Time series data

As we sought to characterize patterns of co-activation, time series data were initially processed so as to remove the estimated contributions of the paradigm. To do so, we ran multiple linear regression (with AFNI's 3dDeconvolve) on the preprocessed functional data with the goal of estimating the residual time series after the inclusion of the following regressors: proximity, velocity, velocity $\times$ proximity, and visual

motion. The regressors were determined based on the circle positions on the screen. Proximity was defined as the Euclidean distance between the two circles. Velocity was the discrete temporal difference of proximity. The visual motion regressor was defined as the velocity tangential to the difference vector of the combined circle-to-circle stimulus, and was calculated by multiplying the angular velocity of the difference vector by the proximity (and accounted for motion energy orthogonal to the relative motion between the circles).

For each run, the regressors were obtained by first decimating the 20 Hz sample rate of stimuli information (used to compute circle paths) to the TR sample rate (1 Hz). Within each run, the regressors were mean-centered to reduce collinearity between simple effects (proximity, velocity) and the velocity $\times$ proximity interaction term, and convolved with a standard hemodynamic response based on the gamma-variate model (Cohen, 1997). In addition, we regressed out any potential block-sustained activation by including a regressor representing the blocks (60-second duration), which was convolved with the standard hemodynamic response. Other regressors included in the model comprised 6 motion parameters (3 linear displacements and 3 angular rotations), and their discrete temporal derivatives. Additionally, to model baseline and drifts of the MR signal, regressors corresponding to polynomial terms up to 4th order were included (for each run separately). To minimize the shock effect, data points in a 15-sec window after shock delivery were discarded from all analyses. Finally, the residual time series for each run were z-scored separately.

The residual time series as defined above was used for the intersubject network

analysis, whose main goal was to characterize networks during approach and retreat conditions, and contrast them. As specified above, approach and retreat varied dynamically throughout the blocks. Therefore, we employed a windowing procedure to extract data segments corresponding to the conditions of interest. Intuitively, the windowing allowed us to select segments of the time series associated with each condition and concatenate them across all runs, generating a final concatenated times series for each condition. First, for each block, the first 15 time points (15 seconds) were discarded, to minimize contributions from block onset. Segment type was determined by considering the velocity regressors (which determined approach vs. retreat). Specifically, transitions in the sign of the velocity regressor indicated the start of a segment type (approach or retreat), to which a 5-second lag was added to account for hemodynamics. Furthermore, based on the 75% cutoff described earlier, data were discarded based on proximity data; in other words, we only considered data in the "75% near space." Finally, all the time points across all blocks and runs assigned to each condition were concatenated for that condition, and constituted the time series data for the condition.

5.2.3 Exploratory and test sets

The total dataset was subdivided into "exploratory" and "test" sets. The idea was to use the exploratory set to fix specific processing choices; with the entire processing pipeline fixed, statistical testing was then applied to separate data in the test set. The size of the exploratory set (N=37) was determined arbitrarily and based on

splitting the data available at a certain date during the data acquisition process; as scanning continued for a bit longer, the test set contained a larger number of participants ($N = 48$). To reiterate, the results reported here are based on the test set alone; thus, processing choices were not optimized or tuned to the test sample, by design.

5.2.4 Regions of interest

We investigated three networks widely studied in the literature: salience, executive, and task negative. From these networks, we employed cortical ROIs (Table 5.1; defined as 5-mm radius spheres) based on the center coordinates provided by previous studies: salience network (Hermans et al., 2011)(13 regions), executive network (Seeley et al., 2007) (12 regions), and task-negative network Fox et al. (2005) (12 regions).

Based on our goal of investigating threat/anxiety, we included additional subcortical ROIs (see Table 5.2 for list of all subcortical regions): amygdala, hippocampus, cerebellum, PAG, habenula, and BNST. For the amygdala, we considered the subregions defined by the Nacewicz et al. (Nacewicz et al., 2014) parcellation, specifically: lateral amygdala; basolateral/medial amygdala; cortical nucleus plus amygdalo-hippocampal area; central plus medial nuclei. For the hippocampus, we focused on its anterior portion because the rodent literature has implicated the ventral hippocampus (believed to correspond to the anterior part in humans) in anxiety-related processing (Bannerman et al., 2004). The hippocampus ROI was

defined by using the hippocampus mask from FreeSurfer and cutting it at the y = +21 plane (MNI coordinates). For the cerebellum, a meta-analysis indicated that Lobules I-IV and Crus II were involved in emotion-related processing (Riedel et al., 2015). Masks for these regions were obtained from the cerebellum parcellation available in FSL (Diedrichsen et al., 2009) (called VIIa Crus II region in the FSL atlas). For the PAG, we modified the mask by (Roy et al., 2014), which was dilated by 1 voxel; in addition, we manually removed the voxels that extended above/below the superior/inferior limits of the original ROI, and those overlapping cerebrospinal fluid. The habenula has been implicated in emotional/motivational processing (Hikosaka, 2010), and here we employed a mask defined according to the Morel atlas, as defined in (Krauth et al., 2010). For the BNST, we employed a recently developed mask based on 7 Tesla data but defined having in mind 3 Tesla data (Theiss et al., 2017).

If two cortical or subcortical ROIs abutted each other, each mask was eroded by 1 voxel from the touching boundary to minimize any potential data "spill over." The exceptions to this rule were the amygdala subregions, BNST, and habenula, because of their very small volume.

Based on exploratory set results, for test set inferences, we focused on the following subcortical ROIs: right lateral amygdala; right PAG, right habenula, left cerebellum crus and right BNST.

Table 5.1: **List of Cortical Regions of interest.** Cortical ROIs were defined via 5-mm radius spheres centered on MNI coordinates provided below.

ROI names	**Coordinates (MNI)**		
	x	**y**	**z**
Salience			
1) frontoinsular cortex L	-34	18	4
2) frontoinsular cortex R	34	22	4
3) Dorsal anterior cingulate cortex	2	10	40
4) Temporo-parietal junction L	-62	-26	36
5) Temporo-parietal junction R	62	-26	36
6) Inferotemporal cortex L	-54	-62	-4
7) Inferotemporal cortex R	54	-54	-8
8) Precentral L	-26	-6	64
9) Precentral R	26	-2	64
10) Dorsolateral prefrontal cortex L	-38	42	24
11) Dorsolateral prefrontal cortex R	34	46	28
12) Inferior frontal gyrus L	-54	6	20
13) Inferior frontal gyrus R	54	10	12
Executive			
14) Orbital frontoinsula L	-36	24	-10
15) Dorsolateral prefrontal cortex R	46	46	14
16) Dorsolateral prefrontal cortex L	-34	46	6
17) Ventrolateral prefrontal cortex R	34	56	-6
18) Ventrolateral prefrontal cortex L	-32	54	-4
19) Frontal operculum R	56	14	14
20) dorsolateral prefrontal cortex / frontal eye field R	30	12	60
21) dorsolateral prefrontal cortex / frontal eye field L	-32	18	50
22) dorsomedial prefrontal cortex	0	36	46
23) Lateral parietal R	38	-56	44
24) Lateral parietal L	-48	-48	48
25) Inferior temporal R	58	-54	-16
Task negative			
26) posterior cingulateprecuneus (PCC)	-3	-38	38
27) Retro-splenial	2	-52	9
28) lateral parietal cortex (LP) L	-50	-64	38
29) lateral parietal cortex (LP) R	50	-64	38
30) medial prefrontal cortex (MPF) L	-4	42	-9
31) medial prefrontal cortex (MPF) R	0	59	16
32) Superior frontal L	-16	44	51
33) Superior frontal R	17	43	51
34) Inferior temporal L	-62	-33	-20
35) Inferior temporal R	66	-18	-20
36) Parahippocampal gyrus L	-22	-26	-20
37) Parahippocampal gyrus R	25	-26	-18

Table 5.2: **List of Subcortical Regions of Interest.** Subcortical ROIs were defined anatomically (see text for details). Numbering is continued from Table 5.1

ROI names
38) Amygdala basolateral/medial R
39) Amygdala basolateral/medial L
40) Amygdala central/medial R
41) Amygdala central/medial L
42) Amygdala Cortical/hippocampal amygdaloid R
43) Amygdala Cortical/hippocampal amygdaloid L
44) Amygdala Lateral R
45) Amygdala Lateral L
46) Hippocampus L
47) Hippocampus R
48) Periaqueductal gray L
49) Periaqueductal gray R
50) Habenula L
51) Habenula R
52) Cerebellum Crus R
53) Cerebellum Crus L
54) Bed nucleus of the stria terminalis L
55) Bed nucleus of the stria terminalis R

5.2.5 Intersubject functional network

In previous intersubject studies, two approaches have been used. First, the time series of an ROI (or voxel) in one subject is correlated with time series data of the same ROI (or voxel) in the remaining subjects (for example, it can be correlated with the average ROI time series data across the "remaining" subjects), yielding a resulting intersubject correlation map (Hasson et al., 2004). Second, intersubject seed-based analysis can be performed, in which the time series of an ROI in one subject is correlated with time series data of a group of ROIs (or voxels) of other subjects (Simony et al., 2016).

A simple, yet powerful extension is to consider intersubject correlations across all pairs of ROIs, which allows the application of the technique to networks. The

procedure to generate an intersubject network is specified in Algorithm 5.1. For a given ROI, first a subject's s data is held out ($\mathbf{Y}_s$), and the rest of the subjects' time series is averaged ($\bar{\mathbf{Y}}_{-s}$). Then, the Pearson correlation between the left-out data and the corresponding data in the remaining subjects is computed: $corr(\mathbf{Y}_s, \bar{\mathbf{Y}}_{-s})$. This basic operation is repeated for all pairs of ROIs to compute an intersubject network for the held-out subject ($\mathbf{ISN}_s$). (For compactness, all of the ROI time series are stacked into a matrix $\mathbf{Y}$ in the Algorithm 5.1) Thus, the $\mathbf{ISN}_s$ is an $N \times N$ matrix, where N is the number of ROIs, and the ij-th matrix element contains the correlation coefficient between the i-th ROI time series of the held-out subject and the average of time series of the j-th ROI of the remaining subjects.

This procedure is then repeated for all subjects. A group matrix ($\mathbf{ISN}_{\mathrm{G}}$) is then obtained by averaging across all subjects. Note that the resulting intersubject network is not necessarily symmetric, because, for each ROI, the time series in the held-out subject ($\mathbf{Y}_s$) is not necessarily equal to the average of all other subjects' time series ($\bar{\mathbf{Y}}_{-s}$). While the $\mathbf{ISN}_{\mathrm{G}}$ matrix in practice will be near symmetric, a simple and intuitive way to mathematically accomplish symmetry is to average the group-level intersubject network with its transpose (where row and column indexes are flipped), leading to a final symmetric group matrix. Finally, the procedures above were performed, separately, for the approach and retreat conditions generating the matrices $\mathbf{ISN}_{\mathrm{G,APPROACH}}$ and $\mathbf{ISN}_{\mathrm{G,RETREAT}}$. Finally, note that in our method the diagonal is also computed because data at a given diagonal entry (i, i) is computed across different brains.

Algorithm 5.1 Computation of group intersubject network

1: **for** $s = 1, 2, ..., N$ **do** // s indicates subject's index

2: $\mathbf{Y}_s \stackrel{\text{def}}{=}$ ROIs' time series of subject s

3: $\bar{\mathbf{Y}}_{-s} = \frac{1}{N-1} \sum_{\substack{i=1 \\ i \neq s}}^{N} \mathbf{Y}_i$ // Average of timeseries of all subjects except s

4: $\mathbf{ISN}_s = corr(\mathbf{Y}_s, \bar{\mathbf{Y}}_{-s})$ // subject level ISN

5: $\mathbf{ISN}_G = \frac{1}{N} \sum_{s=1}^{N} \mathbf{ISN}_s$ // group level ISN

6: $\mathbf{ISN}_G = (\mathbf{ISN}_G + \mathbf{ISN}_G^T)/2$ // symmetrized group level ISN; T: transpose

5.2.6 Dynamic intersubject networks

Considering all data points simultaneously, as done thus far, gives us a static understanding of intersubject networks. Here, we sought to investigate dynamic aspects of network organization. To do so, we computed intersubject networks at each time t and investigated how network properties changed across segments of approach and retreat. Thus, for each segment type (approach and retreat, separately), we considered all of the $t = 0$ time points, then $t = 1$ time points, and so on, separately. The goal was to generate a time series of data at $t = 0$ by concatenating all of the $t = 0$ data across segments. To do so, we concatenated the points $t = k$ (for a fixed k) across segments (Figure 5.3). To account for the hemodynamic delay, we discarded the first 5 seconds of each segment. In this manner, we determined $\mathbf{ISN}_{t=0}$, $\mathbf{ISN}_{t=1}$, and so on. We considered intersubject networks for at $t = 0, 1, ..., 6$ for the approach condition (circles moving closer to each other) and $t = 0, 1, ..., 5$ for the retreat condition (circles are moving away from each other). This was done such that at least 20 data points were available per condition and time "slices" (note

that less data was available for the retreat condition, as some points were discarded following shock events).

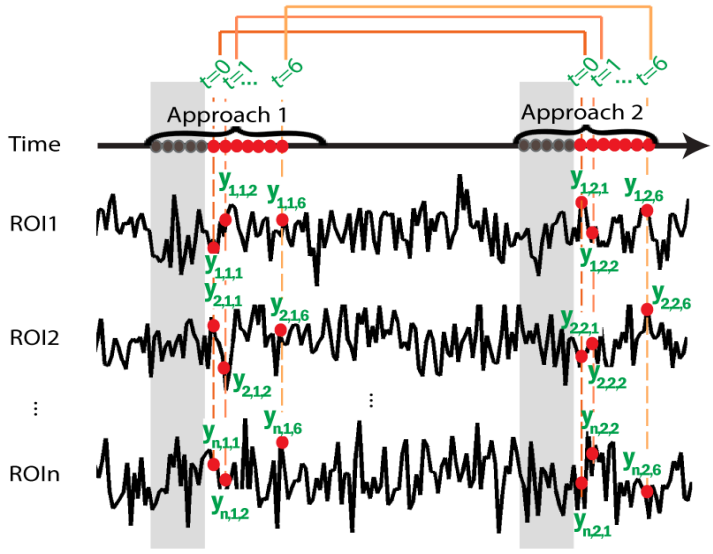

Figure 5.3: **Procedure for data concatenation used in evaluating intersubject network dynamics. Procedure for data concatenation used in evaluating intersubject network dynamics.** Data point y for each time point t was used to compute intersubject networks for approach and retreat, separately. Data point indexes: ROI, condition segment (two approach data segments are illustrated), and time within segment. Briefly, time was used to "slice" and "concatenate" through the ROI time series. Thus, we generated a data time series at t=0 by concatenating all of the t=0 sample (across same-condition segments across blocks and runs), did the same for t=1, and so on. To account for the hemodynamic delay, we discarded the first 5 seconds of each segment (gray part). The resulting data per ROI, time point, and condition, was then investigated in terms of dynamic properties. Time series data were simulated for illustration.

5.2.7 Within- and between-network cohesion

To measure the strength/cohesion of connections within and between networks, we utilized within- and between-network degree, respectively. The cohesion between

network N_1 and network N_2, C_{N_1,N_2}, was defined as follows:

$$C_{N_1,N_2} = \frac{1}{N_{N_1} N_{N_2}} \sum_{i \in N_1, j \in N_2} A_{ij} \tag{5.3}$$

where N_{N_1} is the number of ROIs in network N_1, and N_{N_2} is the number of ROIs in network N_2. A_{ij} is the $\mathbf{ISN}_{\mathrm{G}}$ value between i-th ROI and j-th ROI when i-th ROI belongs to network N_1 and j-th ROI belongs to network N_2. If the above formula, N_1 and N_2 are the same network, then above formula calculates within-network cohesion. In this case both i-th and j-th ROIs belong to the same network.

Defining network cohesion in terms of degree had two advantages. First, we considered both positive and negative weights, unlike most approaches that discard negative weights because many network measures (such as *efficiency*) do not easily handle negative values (Newman, 2010). Most network measures also do not handle self-connections (Newman, 2010), which in standard analysis are not informative anyway ($A_{ii} = 1$). Here, we considered functional connections between the same region (across brains), which could be incorporated in within-network cohesion by considering the terms A_{ii} in the computation of cohesion, C.

To evaluate functional connections between subcortical regions and the salience network (Figure 5.5), we computed a cohesion index that summed all functional connections between a specific subcortical region and all nodes of the salience network. This was performed for the approach and retreat conditions, separately. To test for cohesion, a one-sample t-test (against zero) was employed. A paired t-test was employed to compare cohesion between approach to retreat conditions.

To study dynamic changes to network cohesion, $\mathbf{ISN}_{\mathrm{G}}$ was computed as outlined above at each time t (Figure 5.6). To test for time effects, linear regression was employed, and applied separately to the approach and retreat conditions.

5.3 Results

5.3.1 Intersubject network analysis and statistical approach

Standard network analysis of functional MRI data is based on a adjacency matrix in which each entry is the correlation between time series data for two ROIs within the same participant (Bullmore and Sporns, 2009). Here, we employ a method to extend intersubject correlation analysis to networks. Although we developed this method independently (Najafi and Pessoa, 2016), Simony and colleagues developed essentially the same method and applied it to the study of understanding the task-negative network during narrative comprehension (Simony et al., 2016). We call our version of the method intersubject network analysis. Specifically, the correlations across all pairs of regions is computed (Figure 5.2B), generating a correlation matrix that can be investigated via graph theory techniques (Newman, 2010).

By correlating time series data across participants, intersubject network analysis captures temporal signal properties that are shared by them, deemphasizing fluctuations that are incidental and observed in individual participants (Simony et al., 2016). Another important property of intersubject network analysis is that it can consider the correlation of a region with itself. Whereas in standard analysis this is uninformative (a region's correlation with itself is by definition 1), in intersubject

network analysis the correlation is meaningful (and informative) because the time series data come from different brains.

To develop the method and its application to the dynamic threat paradigm without "peeking" into data, we applied it first to a subset of our entire dataset, which we call the "exploratory set" ($N = 37$), and was used to fix particular processing choices. The results described here were obtained in a separate "test set" ($N = 48$) independent from the exploratory set. Our goal was to enhance reproducibility in a research area plagued by the "curse of flexibility." For example, Poldrack and colleagues (Poldrack et al., 2017) recently enumerated 69,120 different workflows for basic functional MRI analysis alone. We advocate the present approach with exploratory and test sets to imaging studies that do not target very specific questions (which we believe are rarer than typically acknowledged), and/or that include novel methodology (as in the present case). Note that, although we refer to our sets as "exploratory" and "test," our goal was not to attempt discovery and replication in a single study, as in genomics, for example. Specifically, the objective of using an exploratory set was to develop the method and to narrow down the brain regions being investigated.

The total experiment included 6 runs, each of which included 6 blocks. In each block, the circles moved on the screen for 60 seconds; blocks were separated by a 15-second blank screen. We investigated functional connectivity of several regions that are challenging to image with functional MRI, including amygdala subnuclei, the BNST, and the habenula. Although great care was taken at co-registration and functional data were not smoothed, we suggest that region labels

be considered "putative" insofar as higher functional resolution would be required for clearer anatomical attribution. See Methods for further information.

5.3.2 Network cohesion

We determined intersubject correlation matrices for the approach and retreat conditions (see Methods), which allowed us to determine within- and between network cohesion (based on node degree) for the two conditions, and to compare cohesion values during approach vs. retreat. Positive cohesion values indicate that correlations between regions within a network or between networks were on average positive; negative cohesion values indicate that they were on average negative. Note that although our measure of cohesion was based on degree, it is not subject to the recent criticism of using degree when estimating node importance (Power et al., 2013), because that was not our goal here. Importantly, by utilizing node degree, we could parsimoniously employ both positive and negative weights, thus providing a better characterization of network "strength."

During approach (Figure 5.4A), positive network cohesion was detected within the salience network (one-sample t test; $t(47)$ =8.91; p =9.45e-12), and between the salience and executive networks (one-sample t test; $t(47)$ =2.8; p =5.92e-3); negative cohesion was detected between the salience and task-negative networks (one-sample t test; $t(47)$ =-5.61; p =9.68e-7) (nodes tended to be negatively correlated). Furthermore, within-network cohesion was positive in the task-negative network (one-sample t test; $t(47)$ =3.42; p =1.26e-3) (Figure 5.4B-D plots the results as

summary "block matrices" for convenience). Interestingly, a similar pattern of results was observed during retreat (in all cases one-sample t tests; within salience: $t(47)$ =6.63; p =2.74e-8; between salience and executive: $t(47)$ =3.14; p =2.89e-3; between salience and task-negative: $t(47)$ =-3.61; p =7.17e-4). Finally, the direct comparison between approach vs. retreat revealed increased cohesion in the salience network during approach vs. retreat (two-sample paired t test; t(47) =2.67; p =1.01e-2).

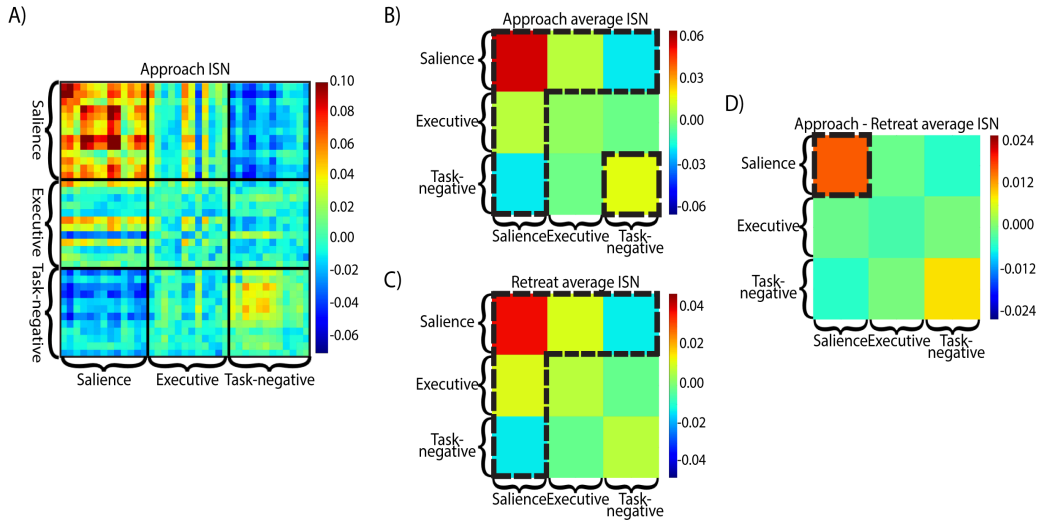

Figure 5.4: **Intersubject group networks. A)** Intersubject network (ISN) during threat approach periods. ROI order corresponds that that of Table 1. **B-D)** Average ISN values for each network at approach (B), retreat (C), and approach minus retreat (D). The dark rectangles surround all of the blocks with significant values (p ¡ 0.01); note that each of the 9 blocks was treated as a unity (the outline extended over multiple of them for diagramming convenience). The color bars indicate differences in intersubject correlation.

In brief, during threat processing (including both approach and retreat periods) both the salience network and its interactions with other networks exhibited the most conspicuous patterns of correlation. In addition, the task-negative network

also exhibited positive cohesion during approach.

5.3.3 Subcortical regions.

Several subcortical regions are involved in threat processing. Furthermore, it has been suggested that some subcortical regions are functionally linked with the salience network under threat conditions (Hermans et al., 2011). Based on existing literature, we considered the amygdala (subdivided into subregions), anterior hippocampus, periaqueductal gray (PAG), and BNST. Based on analysis using the exploratory set, we also report test-set results on the habenula and the cerebellum crus.

Figure 5.5 displays intersubject functional correlations between the salience network and subcortical regions. To evaluate functional connections between subcortical regions and the salience network, we computed a cohesion index that treated the subcortical region as a unit (thus summing degree across all nodes of the salience network). For each of reference, statistical values are provided for each subcortical region and condition in Figure 5.5 (see Methods). During approach, several subcortical regions were positively correlated with the salience network, including the right lateral amygdala, right/left PAG, right/left habenula, left Cerebellum crus, and right BNST. As the pattern was somewhat similar during retreat, only the following regions exhibited robust difference for approach relative to retreat: right basolateral/medial amygdala, right lateral amygdala, left PAG, and right habenula.

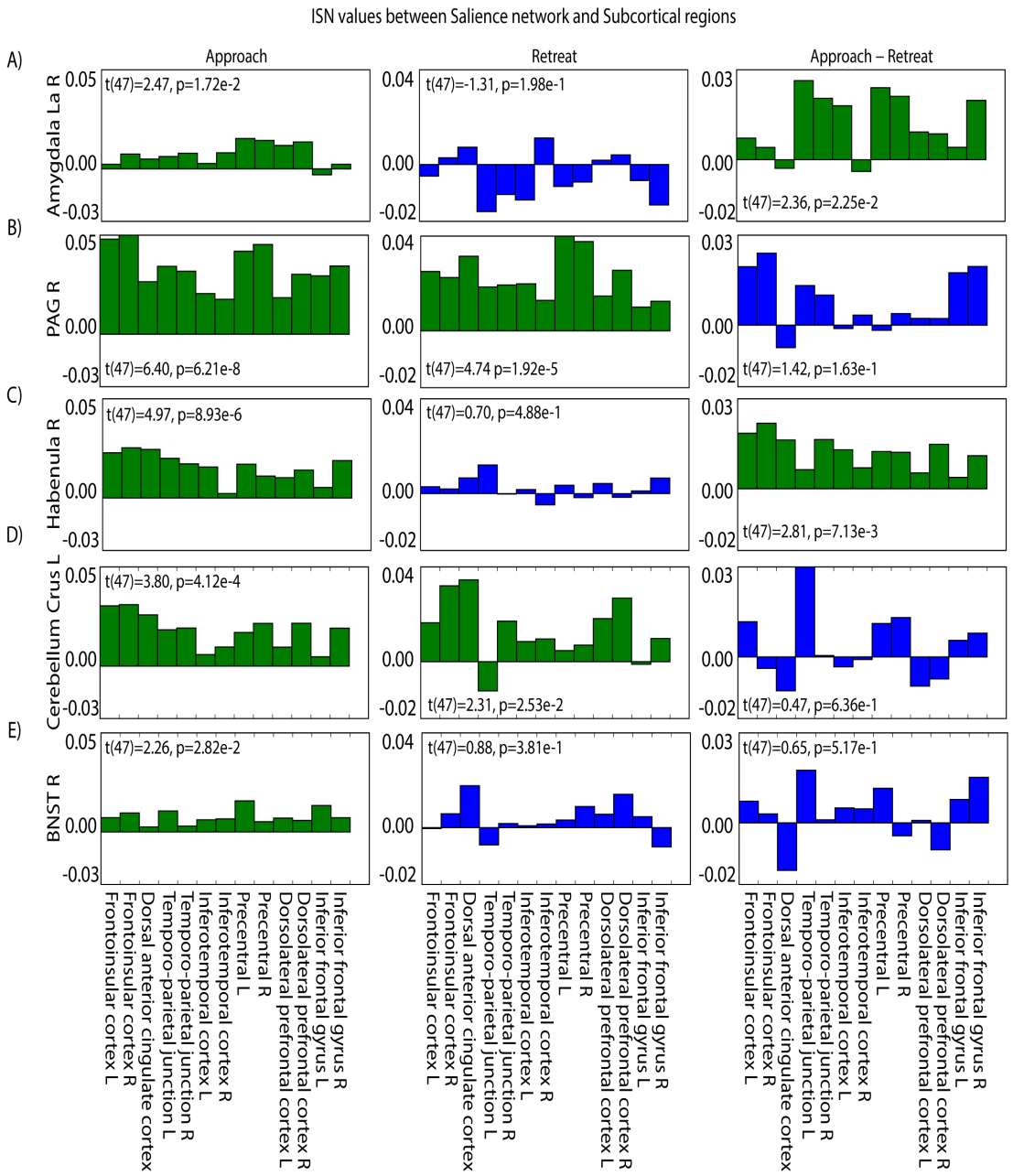

Figure 5.5: **Intersubject functional connections between the salience network and subcortical regions. A)** Amygdala La R; **B)** PAG R; **C)** Habenula R; **D)** Cerebellum Crus L. **E)** BNST R. L/R: left, right; Amygdala La: lateral amygdala, PAG: periaqueductal gray, BNST: Bed nucleus of the stria terminalis. For statistical tests, the entire subcortical region was treated as a unit and cohesion between the region and the salience network was tested. Regions shown with green bar plots were significant (p¡0.05).

5.3.4 Dynamics

As threat level was varied dynamically, we investigated how the intersubject correlation matrix evolved temporally. Figure 5.6 shows the temporal evolution of within- and between-network cohesion during approach and retreat for the salience, executive, and task-negative networks. For the salience network, within-network cohesion increased during approach periods and decreased during retreat periods (for ease of reference, statistical values are provided in the figure). Notice that the reverse was observed for cohesion between the salience and task-negative networks. Overall, all networks exhibited dynamic changes during approach and/or retreat periods (that is, at least one of the slopes was statistically significant).

We also investigated the evolution of the interactions between subcortical regions and the salience network. Although cohesion did not increase robustly during approach periods, decreased cohesion was detected during retreat periods for the left BNST, right habenula, and right PAG (Figure 5.7), revealing that their functional association with the salience network is dynamic.

5.4 Discussion

In the present study, we employed intersubject network analysis to investigate the properties of large-scale networks during threat approach and retreat. A central aim was to investigate the evolution of network properties as threat level varied dynamically. Threat altered network cohesion across the salience, executive, and

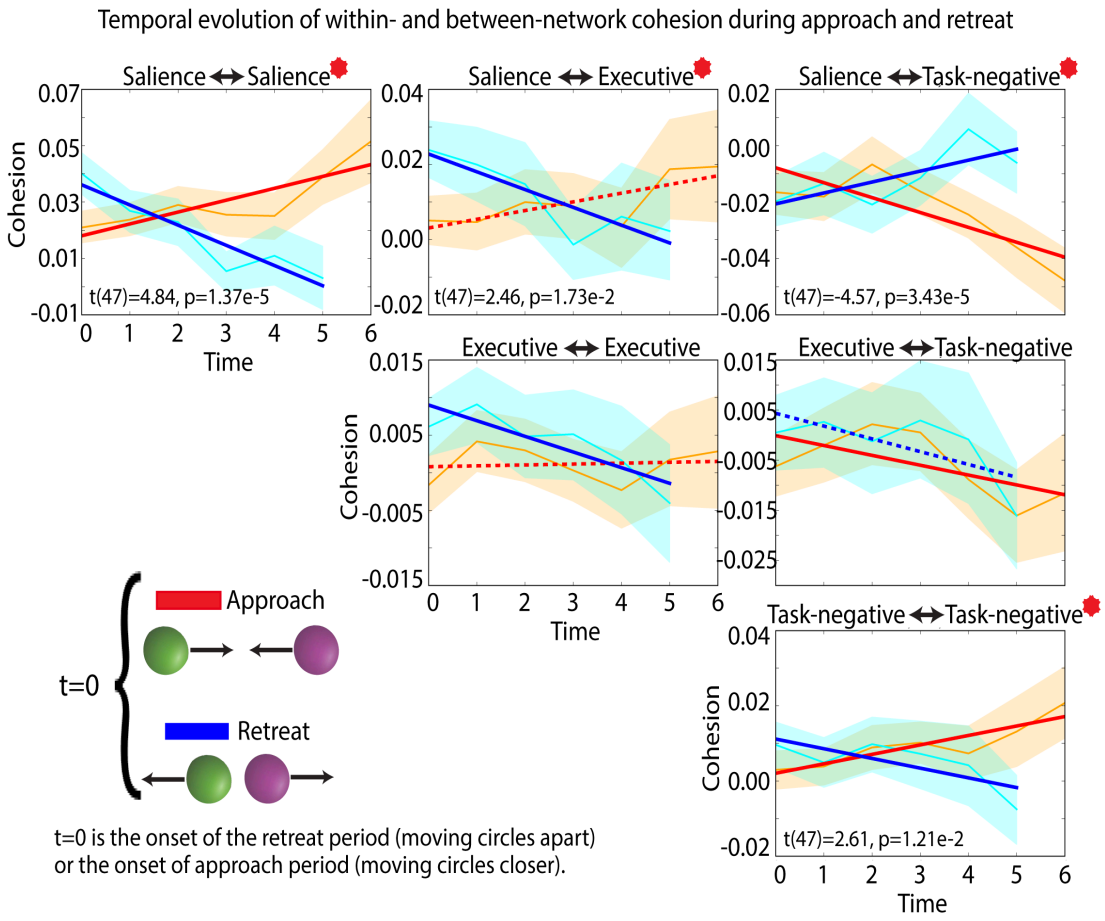

Figure 5.6: **Temporal evolution of cohesion during approach and retreat segments.** Within- and between-network cohesion during approach and retreat for the salience, executive, and task-negative networks. For example, as the circles approach each other, the cohesion within the salience network increased; when the circles retreated, the cohesion within the salience network gradually decreased. The orange line shows cohesion values for approach at different times (with a 90% confidence band); the cyan line shows cohesion values for retreat at different times (90% confidence band). The red/blue line shows the least-squares linear fit to cohesion values during approach/retreat; solid lines indicate fits that were statistically significant (p ¡ 0.05). Time is in seconds (TR = 1 sec); the y-axis shows cohesion (summed degree). The red star indicates that the slope difference was statistically significant (statistical values provided at the bottom left).

task-negative networks, as well as subcortical regions. Importantly, cohesion within and between networks changed dynamically as threat imminence increased and decreased (as circles moved closer and farther to each other). Next, we discuss the

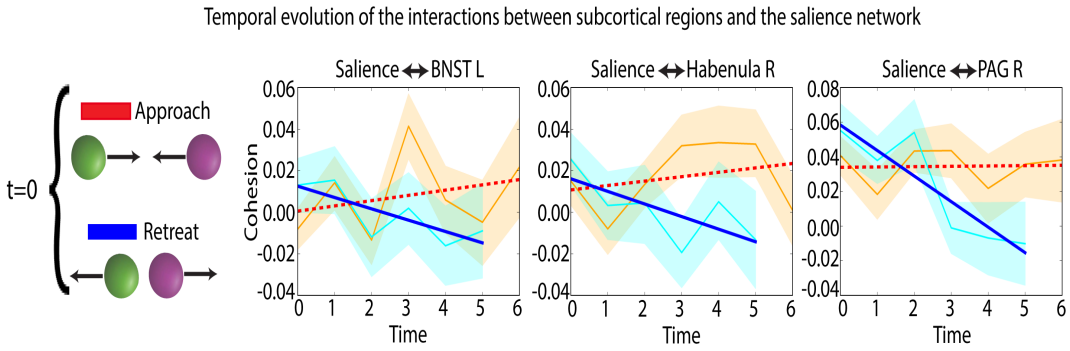

Figure 5.7: **Temporal evolution of cohesion between subcortical regions and the salience network.** Although cohesion did not increase robustly during approach periods, cohesion decreased as a function of time during retreat for the left BNST, right habenula, and right PAG. Conventions as in Figure 5.5.

implications of our main findings.

Standard intersubject correlation analysis has been used to investigate "synchrony" across brains when participants watch the same movie or during other naturalistic conditions, such as hearing extended narratives (Hasson et al., 2004; Nummenmaa et al., 2012, 2014; Stephens et al., 2010). The original formulation was inherently bivariate and considered the same voxel (or region) across participants. The method was recently extended so that a specific voxel/region in one person could be correlated with multiple voxels/regions in other participants (Simony et al., 2016). Independently, we formulated essentially the same method (Najafi and Pessoa, 2016) to perform *intersubject network analysis* with the aim of understanding network organization during dynamic threat. Although we had a specific analysis goal in mind (evaluating network cohesion), more generally, the full range of techniques developed for network analysis can be applied to intersubject data.

Intersubject analyses in general have the advantage that they increase the signal-to-noise ratio by filtering out unwanted contributions to the measured BOLD signal (Simony et al., 2016). This is particularly important for head motion, which can induce significant within-participant correlations (Van Dijk et al., 2012). By computing correlations across participants, the approach essentially eliminates this issue (on the test dataset, head motion parameters exhibited a correlation across subjects of .02).

A central finding of our study was that cohesion within and between networks changed dynamically during periods of approach and retreat. This adds to findings from recent studies that showed how large-scale networks are reorganized during periods of threat (Hermans et al., 2011; McMenamin et al., 2014). Consistent with previous studies, the salience network cohesion increased for approach relative to retreat. Critically, cohesion was not static during approach/retreat segments, but dynamically increased during approach and decreased during retreat. The cohesion between the salience and executive networks followed the same pattern. Notably, the reverse was observed between the salience and task-negative networks. Thus, the salience and task-negative networks cohered more strongly as the circles moved away from each other; movement toward each other made the networks less cohesive. Overall, our findings demonstrate that network cohesion is a dynamic property that depends on threat proximity.

The salience network comprises multiple regions in parietal, frontal, and insular cortices (Menon and Uddin, 2010; Seeley et al., 2007). Sometimes subcortical regions are listed as part of the network, most notably the amygdala and PAG

(Seeley et al., 2007). In the present study, we hypothesized that an extended set of subcortical regions would be closely associated functionally to the salience network. Indeed, this was observed in our data, including, during threat approach, the right lateral amygdala, right/left periaqueductal gray, right/left habenula, cerebellum crus, cerebellum lobes, and right BNST. The present study shows that, in the context of threat processing, these regions are functionally linked to salience processing in paradigms involving threat, and not only during the resting state (see also Hermans et al. (2011)). Importantly, it also shows that this property changes during periods of threat approach relative to retreat, such as in the right lateral amygdala (Figure 5.5).

The findings about the amygdala are particularly noteworthy. Whereas the amygdala is engaged by emotion-laden stimuli and conditions involving acute threat (as in aversive conditioning paradigms), its involvement in potential threat (where threat is not proximal and is relatively uncertain) is less clear (Davis et al., 2010). Some human neuroimaging studies even observed deactivations of the amygdala during conditions of potential threat (Choi et al., 2012; Pruessner et al., 2008; Wager et al., 2009). Here, we saw that multiple amygdala subregions exhibited increased differential functional correlation (approach greater than retreat) with the salience network. These findings are important, because they show that the amygdala is involved during some forms of potential threat (such as the one studied here), as revealed by changes in co-activation patterns (Hermans et al., 2011; McMenamin et al., 2014). It also highlights the need to study functional connectivity and network properties, in addition to evoked responses.

The involvement of the BNST in potential threat was suggested in early work by Davis and colleagues (Davis and Shi, 1999) and has been investigated recently in rodent studies with new neurotechnologies (see Tovote et al. (2015)). Work in humans has revealed the involvement of the BNST in potential threat, too (for reviews, see Fox et al. (2015); Shackman and Fox (2016)). The BNST is rather small and thus challenging to investigate in humans with functional MRI. Nevertheless, recent work at higher resolution and magnetic field strengths (such as 7 Tesla) has been used to generate anatomical masks (Avery et al., 2014; Torrisi et al., 2015), and these appear to be reasonable approximations even at the standard field strength of 3 Tesla (Theiss et al., 2017). An open question concerns the conditions leading to BNST engagement. While some studies suggest that uncertainty may be a major determinant of BNST responses (Alvarez et al., 2011), this is not entirely clear. For example, in a previous study, Mobbs and colleagues (Mobbs et al., 2010) found greater BNST responses for threat approach vs. retreat (although the authors only employed a single level of approach vs. retreat "level," and the activation pattern was very diffuse, thus hard to attribute to the BNST with more confidence). In the present study, the right BNST was more strongly coupled with the salience network during approach (but no differential functional connectivity was detected when comparing approach vs. retreat). Finally, the PAG is another important brain region involved in threat processing (Bandler and Shipley, 1994). Here, we detected increased functional correlations between the right/left PAG with the salience network during approach; we only detected the left PAG as more strongly connected during approach vs. retreat. Note also that the interactions between several subcortical

structures (left BNST, right PAG, and right habenula), and the salience network exhibited temporal properties and decreasing cohesion was observed as the circles moved away from each other during retreat.

In conclusion, in the present study, we employed intersubject network analysis, which allows the investigation of network-level properties "across brains." Our results demonstrate the potential of characterizing emotional processing at the level of distributed networks, and not simply at the level of evoked responses in specific brain regions.

Chapter 6: Conclusions and future directions

6.1 The Brain: An interwoven network

With the advancement of neuroimaging techniques and network science it becomes apparent that the brain's functional system is composed of networks (or communities), and a region-centric perspective of brain functionality is not valid (Najafi et al., 2016; Pessoa, 2014). To analayze large-scale brain networks, the majority of previous studies focused on estimating disjoint communities during the so-call "resting-state" condition. These studies demonstrated that brain regions group into a small number of stable communities (Balenzuela et al., 2010; Power et al., 2012; Yeo et al., 2011). The perspective of modeling functional networks of the brain as disjoint communities potentially originated from the traditional idea of a one-to-one structure-function mapping (Pessoa, 2014, 2015). Though, this type of functional brain characterization reveals important information about the brain organization, it cannot capture the flexible and task dependent mapping between brain regions and their functions (Najafi et al., 2016).

Based on studies of perception, cognition, emotion, and motivation, we have proposed that brain networks are highly interdigitated. Conceptually, this view of brain networks stems from the argument that the mapping from structure to

function is not one-to-one. Instead, the mapping is many-to-many, such that a brain region participates in many functions and similar functions are carried out by many regions (Najafi et al., 2016; Pessoa, 2013, 2014). Consider, briefly, the case of the amygdala. Even a simplified view of its anatomical connectivity shows that, minimally, it belongs to three networks. The first is a "visual network," as the amygdala receives fibers from anterior parts of temporal cortex, and influences visual processing via a set of projections that reach most of ventral occipito-temporal cortex. The second is the well-known "autonomic network," and via connections with the hypothalamus and periaqueductal gray (among many others), the amygdala participates in the coordination of many complex autonomic mechanisms. The third is a "value network," as evidenced by its connectivity with orbitofrontal cortex and medial prefrontal cortex. Thus, the amygdala affiliates with different sets of regions ("communities") in a highly flexible and context-dependent manner (Najafi et al., 2016).

These ideas are related to the "flexible hub theory" by Cole et al. (2013). One component of this framework predicts that "some brain regions flexibly shift their functional connectivity patterns with multiple brain networks across a wide variety of tasks" (Cole et al. (2013); p. 1348). Cole and colleagues suggest that fronto-parietal regions are particularly important "flexible hubs." They described a pattern of functional connectivity that was "representational," where brain-wide functional connectivity patterns across a fronto-parietal community across 64 task states reflected the similarity relationships between tasks, and could be used to identify task states. Work that emphasizes the important role of regions that si-

multaneously participate in multiple networks is an important step in the direction of a richer description of brain network. Our findings in Chapters 3 and 4 are in agreement with their framework, but suggest a dense overlap and highly interwoven organization that extends beyond fronto-parietal regions .

Given the discussion in the previous paragraphs and Chapters 3 and 4, it is instructive to discuss the concept of *modularity* per se. Modularity is a term with multiple connotations in cognitive, brain, and network sciences (for example, see Shallice (1988); Shallice and Cooper (2012)). Although we cannot provide a full account of the issues here (but see Pessoa (2013), Chapter 8; Pessoa (2014)), we briefly comment on the relationship between the presence of overlapping communities and modular structure. In particular, the presence of some overlapping organization in itself does not necessarily imply strong non-modular structure. For instance, as discussed, some nodes may be re-used across communities, particularly brain regions that act as bridges. More generally, systems (neural or otherwise) admit to different degrees of modularity insofar as their components admit to different degrees of isolability (Bechtel and Richardson, 1993). But we suggest that the distribution of network membership values revealed by our analysis, with the associated dense community overlap, reveals a substantial amount of non-modularity in large-scale brain networks at rest and task states (Najafi et al., 2016).

In conclusion, our investigations in Chapters 3 and 4 were driven by the idea that large-scale brain networks will benefit from an overlapping characterization. To detect the overlapping communities, we employed an algorithm based on stochastic variational inference in the mixed-membership stochastic blockmodel by Gopalan

and Blei (2013), which allows each brain region to belong to multiple communities with varying membership strengths. Furthermore, we developed the *membership diversity* based on Shannon entropy (Shannon, 2001), and extended the concept of "modularity" for mixed-membership algorithms to characterize the finer structure and properties of the overlapping communities. Our analysis of rest and task data contribute to understanding several properties of functional brain networks: 1) Overlapping brain networks exhibited general features that resemble those of standard disjoint clustering. However, community membership values spanned the whole range, from weak (closer to 0) to strong (closer to 1), showing that disjoint clustering discretizes important information regarding the association of brain regions to multiple networks. Thus, disjoint communities do not capture the information that is present in mixed communities. 2) Functional diversity of brain regions (that is, the range of functions they participate in) was linearly associated with membership diversity (that is, the extent to which a brain region participates across multiple networks). 3) Task performance substantially altered the structure of functional connectivity across brain regions, and 4) Communicability was enhanced across the brain (that is, modularity decreased during tasks relative to rest). 5) We were also able to study the distribution of "bridge" nodes, including bottleneck and hub bridges.

We conclude that overlapping network methods provide a promising framework to investigate the structure of large-scale brain networks during both rest and tasks states, and task performance altered the role of regions in important ways. Next, we focused on dynamic intersubject networks during anxious anticipation.

6.2 The emotional brain: from a dynamic large-scale network perspective

How do large-scale brain networks reorganize during the waxing and waning of anxious anticipation? The emotional brain has been studied extensively by focusing on a few cortical and subcortical structures, including medial prefrontal cortex, insula, amygdala, and so on. In Chapter 5 by developing dynamic intersubject network technique, we illustrated that characterizing emotional processing from the large-scale brain network perspective has significant benefits for understanding the emotional brain.

Previous large-scale brain network studies are based on the standard network analysis of fMRI data, in which the adjacency matrix is defined by computing the pairwise correlation between brain regions' timeseries data within the same participant (Bullmore and Sporns, 2009). The introduced intersubject network (ISN) method in this dissertation (Chapter 5; Figure 5.1) is a simple yet powerful extension of intersubject correlation (ISC) analysis (Hasson et al., 2004) to networks. In ISC, the time series of a brain region in one subject is correlated with time series data of the same region in the remaining subjects (see Figure 5.2(A1-A2)), whereas in ISN the correlations across all pairs of regions are computed (Figure 5.2(B1-B2)), which results in generating a correlation matrix (see Figure 5.2(B3)), instead of the correlation vector in ISC (see Figure 5.2(A3)). This allowed us to quantify network properties of the emotional brain via graph theory techniques (Newman,

2010). Another advantage of intersubject network analysis is it captures temporal cognitive signal properties that are shared across participants, and deemphasizing non-cognitive fluctuation components, such as head motion effect, in individual participants. Although we developed ISN technique independently (Najafi and Pessoa, 2016), Simony and colleagues developed similar approach to study the task-negative network during narrative comprehension (Simony et al., 2016). Furthermore, we developed dynamic intersubject network analysis (see section 5.2.6 and Figure 5.3), which allows the investigation of the evolution of the properties of large-scale networks across brains.

Our analysis of anxious anticipation contributes to understanding several properties of emotional brain from a large-scale network perspective: 1) We found that threat altered network cohesion across the salience, executive, and task-negative networks, as well as subcortical regions. For example, cohesion increased within the salience network during approach relative to retreat. 2) Functional connections between several subcortical regions and the salience network also increased during approach vs. retreat. The regions included the PAG, habenula, and amygdala, showing that the latter region is involved under conditions of relatively prolonged and uncertain threat, and not only linked to phasic stimuli. 3) The BNST was functionally linked to regions of the salience network during approach, but we did not detect differential engagement when compared to retreat. 4) Cohesion within and between networks changed dynamically as threat imminence increased or decreased. In particular, salience-network cohesion increased during approach and decreased during retreat.

Taken together, our findings unraveled dynamic properties of large-scale networks while threat levels varied continuously. The results demonstrate the potential of characterizing emotional processing at the level of distributed networks, and not simply in a region-centric fashion. In particular, periods during which anxious anticipation waxes and wanes are paralleled by changes to brain network organization.

6.3 Future directions

The work in this dissertation suggested that overlapping network methods provide a promising framework to investigate the structure of large-scale brain networks during both rest and tasks states (Chapters 2 and 3). Also, our results in Chapter 5 illustrated the benefits of investigating the emotional brain at the level of large-scale networks, and not simply at the level of evoked responses in specific brain regions. A future direction is to investigate the overlapping structure of the large-scale brain networks during anxious anticipation. Another future direction is to investigate the temporal changes in overlapping structures at both "rest" and "task conditions". This can be done by combining dynamic intersubject network analysis (Chapter 5) and mixed-membership algorithm (Chapter 3). However, as stated in Chapter 3 (for details see section 3.4) the employed method to investigate the overlapping structure of the brain had some limitations, such as the arbitrary number of ROIs, the requirement of binary adjacency matrix as input, and the number of communities, all of which can be revisited for future studies.

In Chapter 5, we found that dynamic threat altered network cohesion across

the salience, executive, and task-negative networks, as well as subcortical regions. In order to investigate further the re-occurring co-activation patterns and how the communities evolve in time, a promising future direction is to combine multiscale network analysis (Mucha et al., 2010) with temporal intersubject network analysis. Another interesting future direction can be investigating the relationship between behavioral scores and the dynamics of cohesion; for instance, whether there is any relationship between the alteration of network cohesion and anxiety score of the participants.

Chapter A: Supplemental Materials of Chapter 3

A.1 supplemental materials

A.1.1 Membership diversity and participation coefficient

The participation coefficient is a graph-theoretical measure that captures the distribution of edges of a node across all of the communities in a network, and has been used to characterize the type and distribution of hubs in networks (Guimera and Amaral, 2005a,b; Power et al., 2013). It thus expresses a similar property as the membership diversity investigated here. At one level, the distinction between the two can be viewed as fairly subtle. Note, however, that in the case of the participation coefficient, the communities are typically conceptualized as disjoint, and have more or less clearly defined boundaries. The mixed-membership framework offers the possibility to view communities as intrinsically interwoven, such that there is no need to actually define boundaries. This approach may be particularly interesting in situations where dense overlap is present, and a more graded version of community organization is conceptually advantageous.

A.2 Supplemental Figures

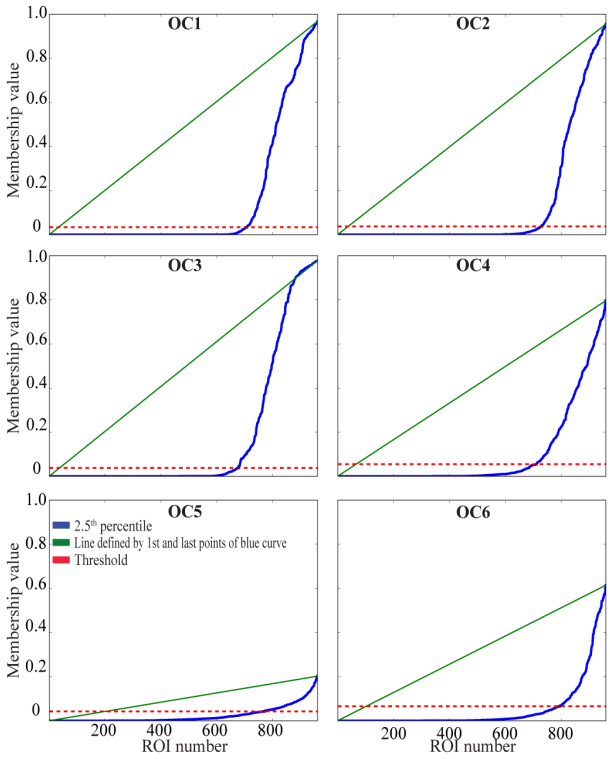

Figure A.1: Membership values for all ROIs reordered in ascending fashion during rest. The elbow was defined as the point at maximal distance from the line passing from the first to the last points of the curve (green line). OC1-OC6: overlapping communities during rest (see Figure 3.6).

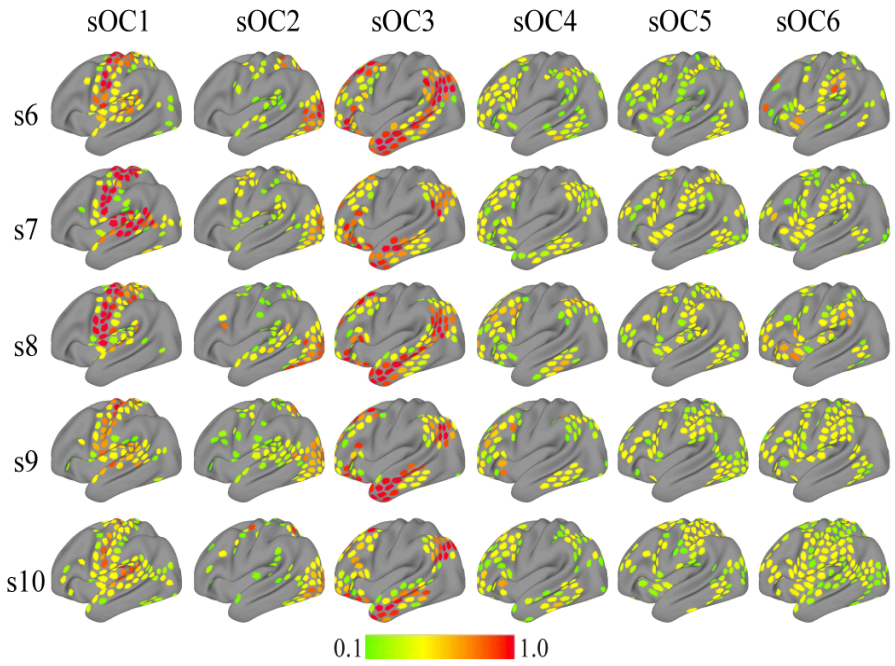

Figure A.2: Overlapping community organization during rest at the individual level for sample participants. Each row depicts the six overlapping communities extracted with the mixed-membership model per subject (sOC). Membership values are thresholded at 0.1 for illustration.

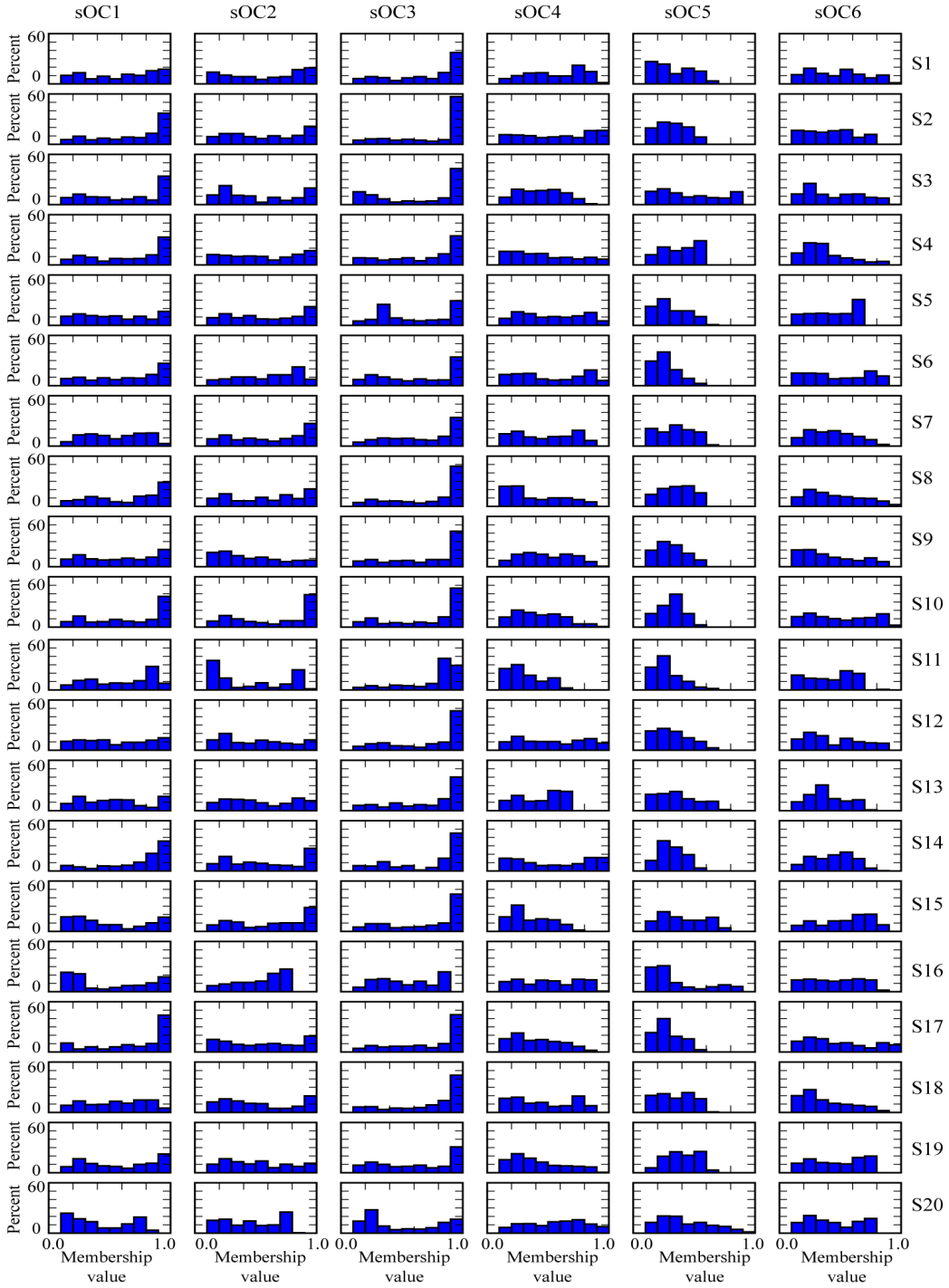

Figure A.3: Histograms of membership values at the individual level (N=94). The results illustrate overlapping community structure at this level of analysis. Membership values were thresholded at 0.1 (thus, the lowest bins are always empty). Subject numbering in the figure does not correspond to the one in Figure A.2.

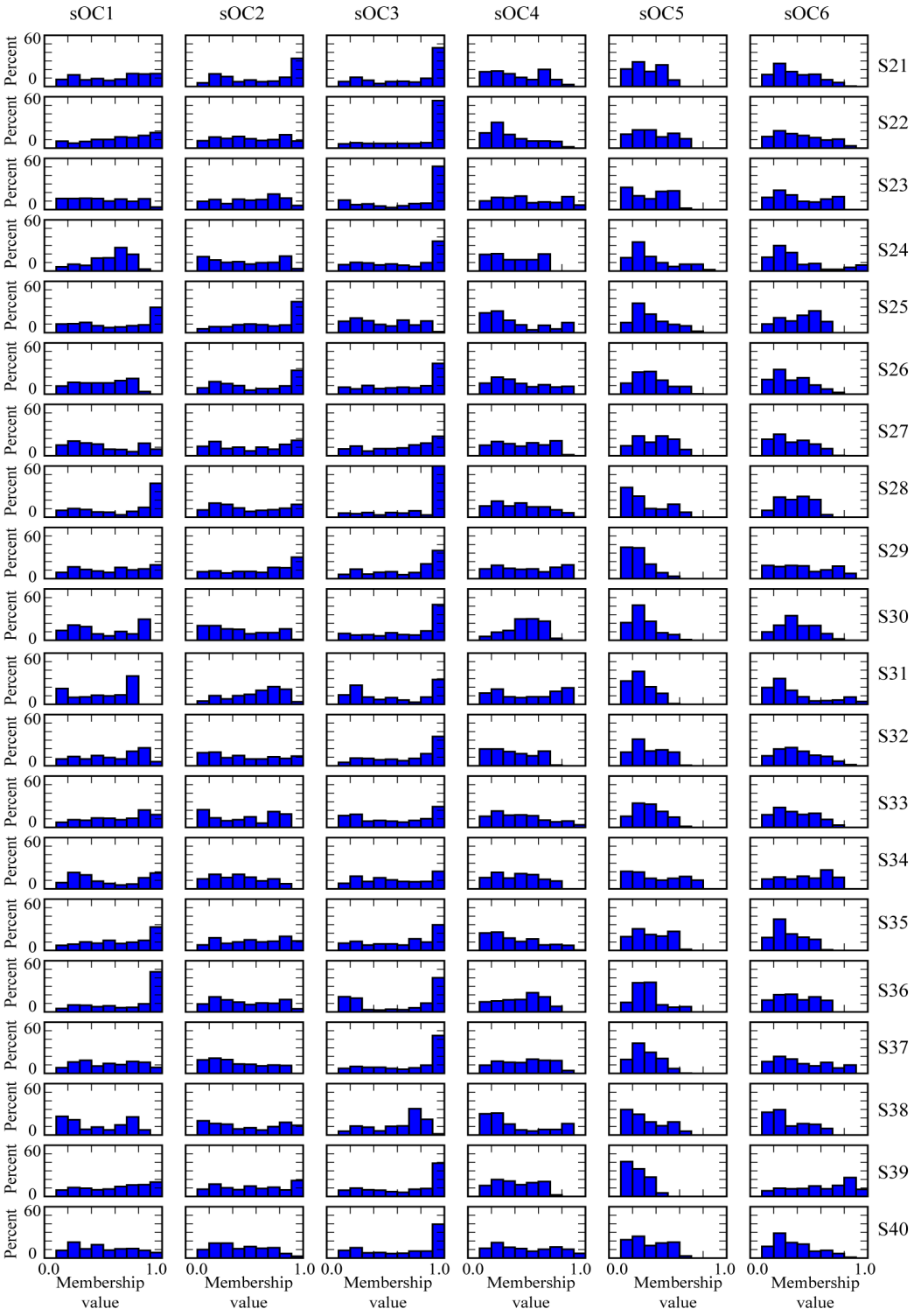

Figure A.4: Figure A.3 Continued

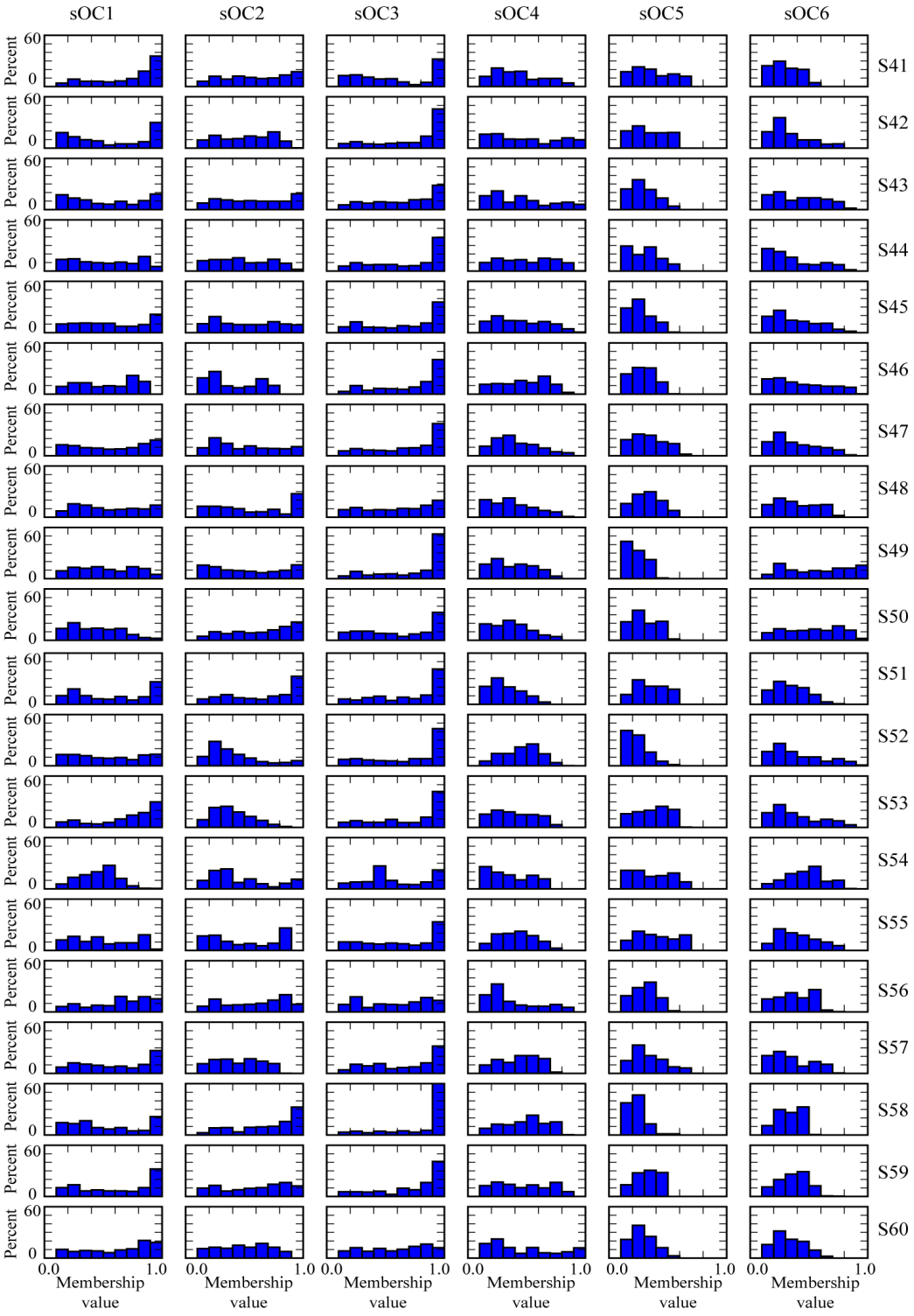

Figure A.5: Figure A.3 Continued

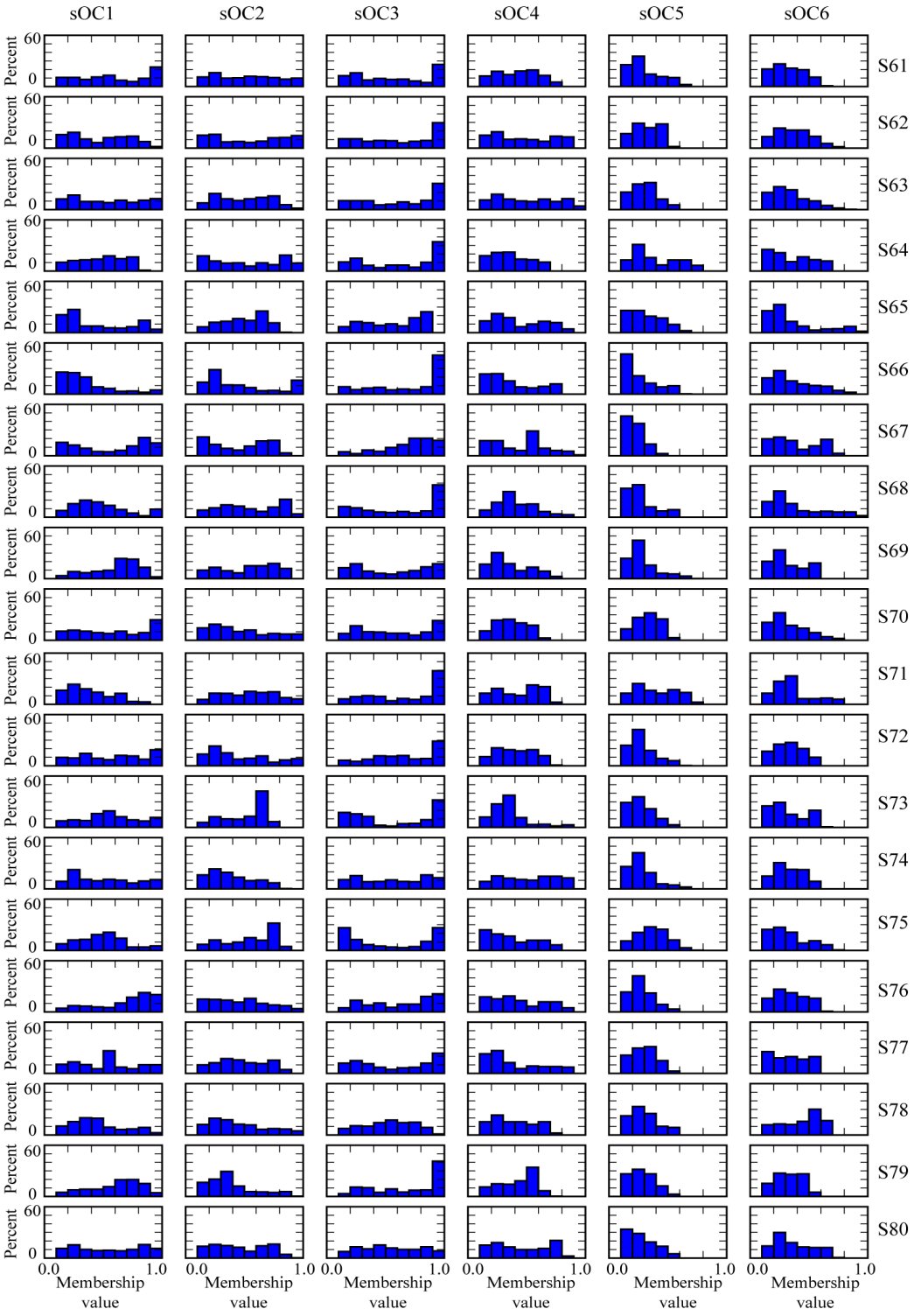

Figure A.6: Figure A.3 Continued

Chapter B: Supplemental Materials of Chapter 4

B.1 supplemental materials

B.1.1 Potential impact of preprocessing pipelines on resting-state and task results

The resting-state data released by the HCP included a preprocessing step called ICA-fixed developed to minimize the effect of non-neural contributions to the time series. However, this preprocessing step was not applied to task data, as these data are typically analyzed via condition contrasts that are assumed to eliminate/minimize potential unwanted signals contributions. In this section, we examine potential effects of the ICA-fixed preprocessing step on overlapping community structure. Moreover, we investigate if the differences in preprocessing between resting-state and task data may have contributed to differences in the results found for the two types of condition.

To reanalyze the resting-state condition, we employed the minimally preprocessed data (provided with the HCP distribution, without the ICA-fixed procedure), and removed potential head-motion contributions via the same procedure utilized for the task conditions. Specifically, this involved regressing out 12 motion param-

eters (6 motion parameters estimated from the rigid-body transformation to the reference image acquired at the beginning of the scan and their temporal derivatives) and censoring the data as described in the Data censoring section of the main text (section 4.2.1.3). We name this resting-state dataset as non-ICA-fixed.

In this manner, we ran the mixed-membership algorithm on non-ICA-fixed resting-state data. The six overlapping communities are depicted in Figure B.1. The results were qualitatively very similar to the ICA-fixed results, as visual inspection indicates (see Figure 3.6). At a quantitative level the median similarity between the corresponding communities was very high (median cosine similarity of 92.24.7%). Thus, community structure at rest was fairly insensitive to the particular preprocessing choice.

Next, we compared overlapping communities at rest and during tasks (as in Section 4.3.1 of the main text), but this time with resting-state results based on non-ICA-fixed preprocessing. The cosine similarity (equation 3.1) between the overlapping communities extracted at rest (non-ICA-fixed) and the working memory and emotion tasks is shown in Figures B.2 and B.3, respectively. Comparison of the results based on non-ICA-fixed (Figures B.2 and B.3) with the results based on ICA-fixed (Figures 4.4 and 4.5) preprocessing reveals that the choice of preprocessing had very small effects. To quantify the differences, we computed the root mean squared error (RMSE) between the corresponding results based on ICA-fixed and non-ICA-fixed preprocessing. Comparison of the matrices in Figure 4.4 and Figure B.2 revealed an RMSE of 0.0683; comparison of the matrices in Figure 4.5 and Figure B.3 revealed an RMSE of 0.0700.

Taken together, the analyses of this section show that the original use of ICA-fixed resting-state data did not have a major impact on the results. Importantly, comparing rest and task data that utilize the same preprocessing steps (non-ICA-fixed) largely reproduced the results reported in the main text.

B.1.2 Potential impact of community size on bridge type

In the investigation of node types, we employed degree as a way to classify bridges; those with low degree were called bottleneck bridges and those with high degree were called hub bridges. One issue that has been investigated in the past is the extent to which degree is potentially confounded with community size (Power et al. 2013). Thus, we evaluated if this potential confound influenced our results, in particular, the number of hub bridges as communities became larger. In other words, if employing degree to label the type of bridge tended to generate a higher number of hub bridges rather than bottleneck bridges as a function of community size. First, we defined two parameters: relative bridge type (RBT), and community size.

*Relative bridge type (RBT).*This parameter determined the percentage of hub bridges relative to bottleneck bridges for each community. As described in the main text, bridgeness scores were continuous. Thus, here we defined high bridgeness scores based on percentile threshold values of 80% to 95% (steps of 5%), so that our analysis would not depend on a single, specific threshold (see Figure 14 for illustration of these cut-off values). Then, each communitys relative bridge type

was computed as

$$\mathrm{RBT}_k = (\#\mathrm{HB}_k - \#\mathrm{BB}_k)/N_k \times 100 \tag{B.1}$$

where k indexed the community, $\#\mathrm{HB}_k$ was the number of hub bridges for community k, $\#\mathrm{BB}_k$ was the number of bottleneck bridges for community k, and N_k was the number of ROIs for community k.

Community size. The community size, N_i, was varied from 50 to 250 ROIs in steps of 20, and was obtained by varying the membership value threshold to obtain the number of target ROIs. Note that in all cases the threshold was higher than the threshold value shown in Figure A.1.

Next, we evaluated the association between the relative proportion of bridge type (RBT) and community size. This was evaluated at each value of RBT threshold (80% to 95%, steps of 5%) and community size, for the three conditions (rest, working memory, and emotion). In each case, 5000 bootstrapping samples were computed. Overall, we found no systematic relationship between RBT and community size. Figure B.4 shows some of the most extreme values (larger relationships) and Figure B.5 shows the entire set of results. Only in a few cases the Spearman correlation between the two variables exceeded 0.25, say, indicating that the variance explained was fairly modest. Importantly, even the direction of the relationship was not consistent across conditions. Based on these analyses, it thus appear that community size has no appreciable effect on our node taxonomy.

B.1.3 Potential impact of number of communities on modularity scores

In section 4.3.3, overlapping modularity was found to be higher during rest compared to both task conditions (Figure 4.10). For that analysis, we fixed the number of communities at $k = 6$ across the three conditions. Although the mixed-membership algorithm does not determine communities based on modularity maximization (other algorithms do; see Fortunato, 2010), it is possible that the overlapping modularity results were affected, at least in part, by choosing the same number of communities across all conditions. This is particularly the case because while $k = 6$ was optimal for the rest condition, the average held-out likelihood was slightly higher for $k = 5$ for working memory, and slightly higher for $k = 4$ for emotion. Therefore, it is possible that the slightly suboptimal number of communities for the task conditions inflated the difference in modularity found for rest relative to tasks.

Here, we examined modularity scores by running the mixed-membership algorithm on working memory data with $k = 5$ and on emotion data with $k = 4$. Modularity scores were determined as in section 4.2.7. Figure B.6 shows the distribution of overlapping modularity scores for the two task conditions, together with those at rest (with $k = 6$). Overlapping modularity was higher at rest (mean and standard deviation: 0.3180.016) than both tasks (working memory: 0.1910.016; emotion: 0.2650.011). Note that modularity scores for the task conditions was higher than when $k = 6$ was used for both tasks (working memory: 0.1830.018; emotion: 0.2110.015; see Figure 4.10). Nevertheless, overlapping modularity at rest

was higher than during task conditions.

B.2 Supplemental Figures

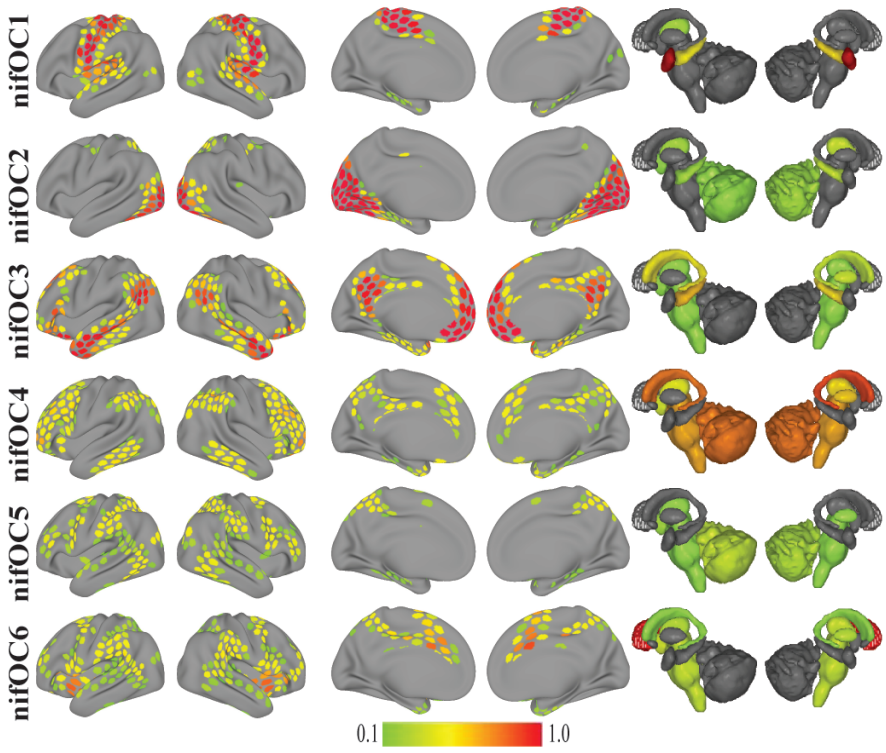

Figure B.1: Overlapping communities detected with non-ICA-fixed preprocessing (see text). The color of cortical and subcortical regions reflects the membership value of each region to each community. Membership values are thresholded at 0.1 for illustration.

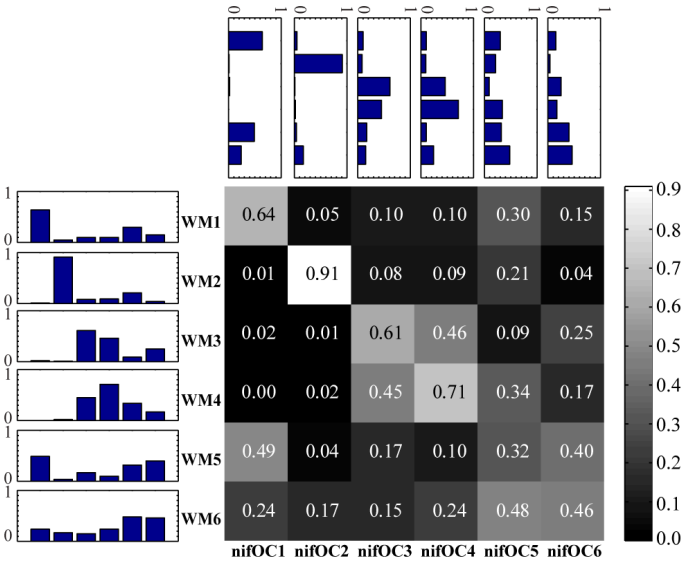

Figure B.2: Cosine similarity between overlapping communities at rest (non-ICA-fixed) and during the working memory task (WM). The matrix displays the cosine similarity between rest and WM membership values.

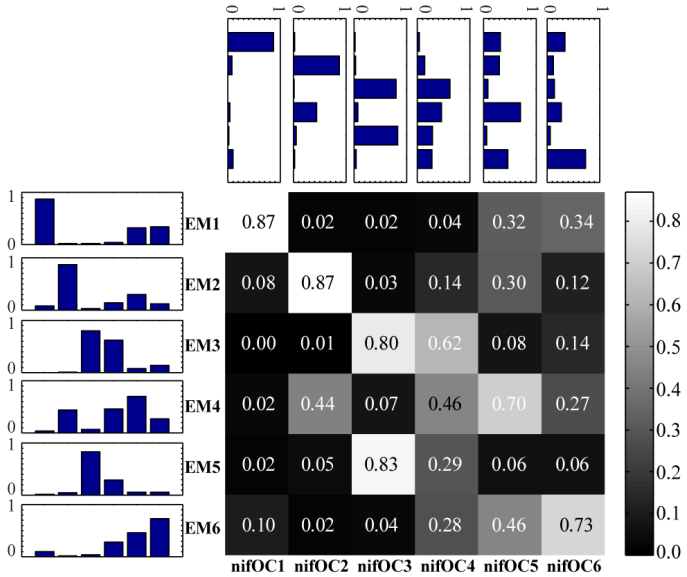

Figure B.3: Cosine similarity between overlapping communities at rest (non-ICA-fixed) and during the emotion task (EM). The matrix displays the cosine similarity between rest and EM membership values.

Figure B.4: Relationship between community size and relative bridge type (RBT). For different community sizes, we determined the RTB index. In these plots, the percentile threshold for high bridgeness score was 95%. For each community size, 5000 bootstrapping samples were computed, and the RBT was computed for each; the mean and the 25-75% ranges are shown. The linear fits are for illustration only.

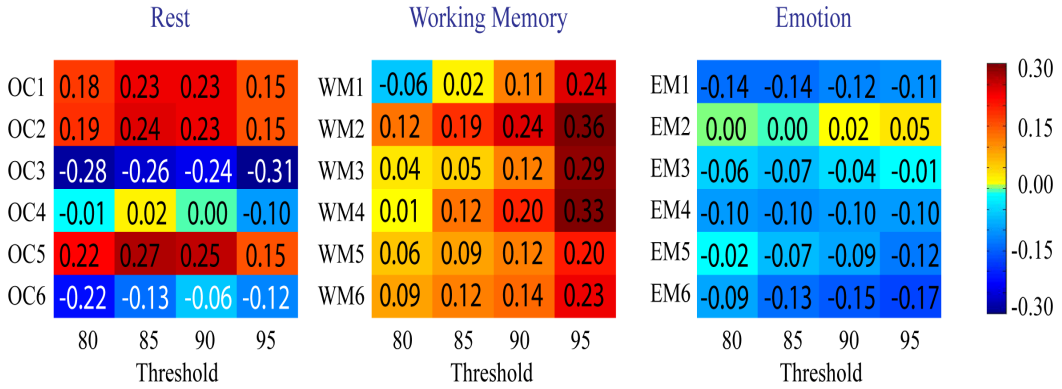

Figure B.5: Summary results for the relationship between community size and relative bridge type (RBT). The same procedure as in Figure B.4 was employed, but here each matrix entry summarizes the relationship between relative bridge type and community size for a given community and bridgeness threshold; thus, for each cell, community size was varied from 50 to 250 (and in each case, 5000 bootstrapping samples were employed). The relationship was assessed via spearmans rank correlation between RBT and community size.

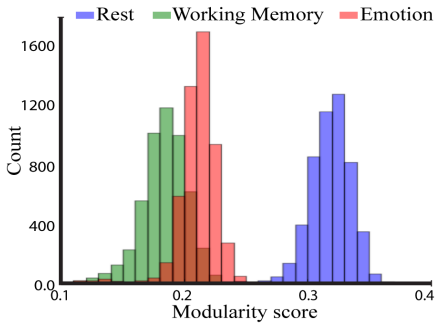

Figure B.6: Modularity scores of overlapping communities during rest ($k = 6$), working memory ($k = 5$) and emotion ($k = 4$). The histograms depict the whole-brain modularity scores across 2,500 iterations (each modularity score was obtained by summing modularity scores across communities).

Bibliography

Airoldi, E. M., Blei, D., Erosheva, E. A. and Fienberg, S. E. (2014), *Handbook of Mixed Membership Models and Their Applications*, Chapman and Hall/CRC.

Airoldi, E. M., Blei, D. M., Fienberg, S. E. and Xing, E. P. (2008), 'Mixed membership stochastic blockmodels', *Journal of Machine Learning Research* **9**(Sep), 1981–2014.

Alexander, G. E., DeLong, M. R. and Strick, P. L. (1986), 'Parallel organization of functionally segregated circuits linking basal ganglia and cortex', *Annual review of neuroscience* **9**(1), 357–381.

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T. and Calhoun, V. D. (2012), 'Tracking whole-brain connectivity dynamics in the resting state', *Cerebral cortex* p. bhs352.

Alvarez, R. P., Chen, G., Bodurka, J., Kaplan, R. and Grillon, C. (2011), 'Phasic and sustained fear in humans elicits distinct patterns of brain activity', *Neuroimage* **55**(1), 389–400.

Amaral, D., Price, J., Pitkanen, A. and Carmichael, S. (n.d.), 'Anatomical organization of the primate amygdaloid complex, aggleton jp, the amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction, 1992, 1-66'.

Ames, D. L., Honey, C. J., Chow, M. A., Todorov, A. and Hasson, U. (2015), 'Contextual alignment of cognitive and neural dynamics', *Journal of cognitive neuroscience* .

Anderson, M. L., Kinnison, J. and Pessoa, L. (2013), 'Describing functional diversity of brain regions and brain networks', *Neuroimage* **73**, 50–58.

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A. and Gee, J. C. (2011), 'A reproducible evaluation of ants similarity metric performance in brain image registration', *Neuroimage* **54**(3), 2033–2044.

Avery, S. N., Clauss, J. A., Winder, D. G., Woodward, N., Heckers, S. and Blackford, J. U. (2014), 'Bnst neurocircuitry in humans', *Neuroimage* **91**, 311–323.

Balenzuela, P., Chernomoretz, A., Fraiman, D., Cifre, I., Sitges, C., Montoya, P. and Chialvo, D. R. (2010), 'Modular organization of brain resting state networks in chronic back pain patients', *Frontiers in neuroinformatics* **4**, 116.

Bandler, R. and Shipley, M. T. (1994), 'Columnar organization in the midbrain periaqueductal gray: modules for emotional expression?', *Trends in neurosciences* **17**(9), 379–389.

Bannerman, D., Rawlins, J., McHugh, S., Deacon, R., Yee, B., Bast, T., Zhang, W.-N., Pothuizen, H. and Feldon, J. (2004), 'Regional dissociations within the hippocampusmemory and anxiety', *Neuroscience & Biobehavioral Reviews* **28**(3), 273–283.

Buckner, R. L., Krienen, F. M. and Yeo, B. T. (2013), 'Opportunities and limitations of intrinsic functional connectivity mri', *Nature neuroscience* **16**(7), 832–837.

Bullmore, E. and Sporns, O. (2009), 'Complex brain networks: graph theoretical analysis of structural and functional systems', *Nature Reviews Neuroscience* **10**(3), 186–198.

Bzdok, D., Laird, A. R., Zilles, K., Fox, P. T. and Eickhoff, S. B. (2013), 'An investigation of the structural, connectional, and functional subspecialization in the human amygdala', *Human brain mapping* **34**(12), 3247–3266.

Chang, C. and Glover, G. H. (2010), 'Time–frequency dynamics of resting-state brain connectivity measured with fmri', *Neuroimage* **50**(1), 81–98.

Chen, D., Shang, M., Lv, Z. and Fu, Y. (2010), 'Detecting overlapping communities of weighted networks via a local algorithm', *Physica A: Statistical Mechanics and its Applications* **389**(19), 4177–4187.

Choi, J. M., Padmala, S. and Pessoa, L. (2012), 'Impact of state anxiety on the interaction between threat monitoring and cognition', *NeuroImage* **59**(2), 1912–1923.

Clauset, A., Newman, M. E. and Moore, C. (2004), 'Finding community structure in very large networks', *Physical review E* **70**(6), 066111.

Cocchi, L., Zalesky, A., Fornito, A. and Mattingley, J. B. (2013), 'Dynamic cooperation and competition between brain systems during cognitive control', *Trends in cognitive sciences* **17**(10), 493–501.

Cohen, M. S. (1997), 'Parametric analysis of fmri data using linear systems methods', *Neuroimage* **6**(2), 93–103.

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S. and Petersen, S. E. (2014), 'Intrinsic and task-evoked network architectures of the human brain', *Neuron* **83**(1), 238–251.

Cole, M. W., Pathak, S. and Schneider, W. (2010), 'Identifying the brain's most globally connected regions', *Neuroimage* **49**(4), 3132–3148.

Cole, M. W., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A. and Braver, T. S. (2013), 'Multi-task connectivity reveals flexible hubs for adaptive task control', *Nature neuroscience* **16**(9), 1348–1355.

Cox, R. W. (1996), 'Afni: software for analysis and visualization of functional magnetic resonance neuroimages', *Computers and Biomedical research* **29**(3), 162–173.

Davis, M. and Shi, C. (1999), 'The extended amygdala: are the central nucleus of the amygdala and the bed nucleus of the stria terminalis differentially involved in fear versus anxiety?', *Annals of the New York Academy of Sciences* **877**(1), 281–291.

Davis, M., Walker, D. L., Miles, L. and Grillon, C. (2010), 'Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety', *Neuropsychopharmacology* **35**(1), 105–135.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E. and Ramnani, N. (2009), 'A probabilistic mr atlas of the human cerebellum', *Neuroimage* **46**(1), 39–46.

Duda, R. O., Hart, P. E. and Stork, D. G. (2012), *Pattern classification*, John Wiley & Sons.

Efron, B. (1979), 'Bootstrap methods: another look at the jackknife', *The annals of Statistics* pp. 1–26.

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., Miller, K. L., Ugurbil, K. and Yacoub, E. (2010), 'Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging', *PloS one* **5**(12), e15710.

Fischl, B. (2012), 'Freesurfer', *Neuroimage* **62**(2), 774–781.

Fortunato, S. (2010), 'Community detection in graphs', *Physics reports* **486**(3), 75–174.

Fox, A. S. and Kalin, N. H. (2014), 'A translational neuroscience approach to understanding the development of social anxiety disorder and its pathophysiology', *American Journal of Psychiatry* **171**(11), 1162–1173.

Fox, A. S., Oler, J. A., Tromp, D. P., Fudge, J. L. and Kalin, N. H. (2015), 'Extending the amygdala in theories of threat processing', *Trends in neurosciences* **38**(5), 319–329.

Fox, E., Russo, R. and Georgiou, G. A. (2005), 'Anxiety modulates the degree of attentive resources required to process emotional faces', *Cognitive, Affective, & Behavioral Neuroscience* **5**(4), 396–404.

Friston, K., Buechel, C., Fink, G., Morris, J., Rolls, E. and Dolan, R. (1997), 'Psychophysiological and modulatory interactions in neuroimaging', *Neuroimage* **6**(3), 218–229.

Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A.-M., Cruciat, C.-M. et al. (2002), 'Functional organization of the yeast proteome by systematic analysis of protein complexes', *Nature* **415**(6868), 141–147.

Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R. et al. (2013), 'The minimal preprocessing pipelines for the human connectome project', *Neuroimage* **80**, 105–124.

Good, B. H., de Montjoye, Y.-A. and Clauset, A. (2010), 'Performance of modularity maximization in practical contexts', *Physical Review E* **81**(4), 046106.

Gopalan, P. K. and Blei, D. M. (2013), 'Efficient discovery of overlapping communities in massive networks', *Proceedings of the National Academy of Sciences* **110**(36), 14534–14539.

Gordon, E. M., Laumann, T. O., Adeyemo, B. and Petersen, S. E. (2015), 'Individual variability of the system-level organization of the human brain', *Cerebral Cortex* p. bhv239.

Goulas, A., Schaefer, A. and Margulies, D. S. (2015), 'The strength of weak connections in the macaque cortico-cortical network', *Brain Structure and Function* **220**(5), 2939–2951.

Greve, D. N. and Fischl, B. (2009), 'Accurate and robust brain image alignment using boundary-based registration', *Neuroimage* **48**(1), 63–72.

Grupe, D. W. and Nitschke, J. B. (2013), 'Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective', *Nature Reviews Neuroscience* **14**(7), 488–501.

Grupe, D. W., Oathes, D. J. and Nitschke, J. B. (2012), 'Dissecting the anticipation of aversion reveals dissociable neural networks', *Cerebral cortex* p. bhs175.

Guimera, R. and Amaral, L. A. N. (2005*a*), 'Cartography of complex networks: modules and universal roles', *Journal of Statistical Mechanics: Theory and Experiment* **2005**(02), P02001.

Guimera, R. and Amaral, L. A. N. (2005*b*), 'Functional cartography of complex metabolic networks', *Nature* **433**(7028), 895–900.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J. and Sporns, O. (2008), 'Mapping the structural core of human cerebral cortex', *PLoS Biol* **6**(7), e159.

Hariri, A. R., Bookheimer, S. Y. and Mazziotta, J. C. (2000), 'Modulating emotional responses: effects of a neocortical network on the limbic system', *Neuroreport* **11**(1), 43–48.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G. and Malach, R. (2004), 'Intersubject synchronization of cortical activity during natural vision', *science* **303**(5664), 1634–1640.

Hermans, E. J., van Marle, H. J., Ossewaarde, L., Henckens, M. J., Qin, S., van Kesteren, M. T., Schoots, V. C., Cousijn, H., Rijpkema, M., Oostenveld, R. et al. (2011), 'Stress-related noradrenergic activity prompts large-scale neural network reconfiguration', *Science* **334**(6059), 1151–1153.

Hikosaka, O. (2010), 'The habenula: from stress evasion to value-based decision-making', *Nature reviews neuroscience* **11**(7), 503–513.

Hilgetag, C.-C., O'Neill, M. A., Young, M. P. et al. (1996), 'Indeterminate organization of the visual system', *SCIENCE-NEW YORK THEN WASHINGTON-* pp. 776–776.

Hoffman, M. D., Blei, D. M., Wang, C. and Paisley, J. W. (2013), 'Stochastic variational inference.', *Journal of Machine Learning Research* **14**(1), 1303–1347.

Huettel, S. A., Song, A. W. and McCarthy, G. (2004), *Functional magnetic resonance imaging*, Vol. 1, Sinauer Associates Sunderland.

Iglesias, J. E., Liu, C.-Y., Thompson, P. M. and Tu, Z. (2011), 'Robust brain extraction across datasets and comparison with publicly available methods', *IEEE transactions on medical imaging* **30**(9), 1617–1634.

Jones, E. (2006), 'The thalamus revisited'.

Karahanoğlu, F. I. and Van De Ville, D. (2015), 'Transient brain activity disentangles fmri resting-state dynamics in terms of spatially and temporally overlapping networks', *Nature communications* **6**.

Kinnison, J., Padmala, S., Choi, J.-M. and Pessoa, L. (2012), 'Network analysis reveals increased integration during emotional and motivational processing', *Journal of Neuroscience* **32**(24), 8361–8372.

Kirby, L. A. and Robinson, J. L. (2015), 'Affective mapping: An activation likelihood estimation (ale) meta-analysis', *Brain and cognition* .

Krauth, A., Blanc, R., Poveda, A., Jeanmonod, D., Morel, A. and Székely, G. (2010), 'A mean three-dimensional atlas of the human thalamus: generation from multiple histological data', *Neuroimage* **49**(3), 2053–2062.

Lahnakoski, J. M., Glerean, E., Jääskeläinen, I. P., Hyönä, J., Hari, R., Sams, M. and Nummenmaa, L. (2014), 'Synchronous brain activity across individuals underlies shared psychological perspectives', *NeuroImage* **100**, 316–324.

Laird, A. R., Lancaster, J. J. and Fox, P. T. (2005), 'Brainmap', *Neuroinformatics* **3**(1), 65–77.

Lancaster, J. L., Laird, A. R., Eickhoff, S. B., Martinez, M. J., Fox, P. M. and Fox, P. T. (2012), 'Automated regional behavioral analysis for human brain images', *Frontiers in neuroinformatics* **6**, 23.

Lancichinetti, A., Fortunato, S. and Kertész, J. (2009), 'Detecting the overlapping and hierarchical community structure in complex networks', *New Journal of Physics* **11**(3), 033015.

Liu, X. and Duyn, J. H. (2013), 'Time-varying functional network information extracted from brief instances of spontaneous brain activity', *Proceedings of the National Academy of Sciences* **110**(11), 4392–4397.

Magurran, A. (2004), 'Measuring biological diversity. blackwells'.

Marcus, D., Harwell, J., Olsen, T., Hodge, M., Glasser, M., Prior, F., Jenkinson, M., Laumann, T., Curtiss, S. and Van Essen, D. (2011), 'Informatics and data mining tools and strategies for the human connectome project', *Frontiers in neuroinformatics* **5**, 4.

McMenamin, B. W., Langeslag, S. J., Sirbu, M., Padmala, S. and Pessoa, L. (2014), 'Network organization unfolds over time during periods of anxious anticipation', *Journal of Neuroscience* **34**(34), 11261–11273.

Menon, V. and Uddin, L. Q. (2010), 'Saliency, switching, attention and control: a network model of insula function', *Brain Structure and Function* **214**(5-6), 655–667.

Mesulam, M.-M. (1998), 'From sensation to cognition.', *Brain* **121**(6), 1013–1052.

Mesulam, M. et al. (1981), 'A cortical network for directed attention and unilateral neglect', *Annals of neurology* **10**(4), 309–325.

Middleton, F. A. and Strick, P. L. (2000), 'Basal ganglia and cerebellar loops: motor and cognitive circuits', *Brain research reviews* **31**(2), 236–250.

Miller, E. K. and Cohen, J. D. (2001), 'An integrative theory of prefrontal cortex function', *Annual review of neuroscience* **24**(1), 167–202.

Mobbs, D., Yu, R., Rowe, J. B., Eich, H., FeldmanHall, O. and Dalgleish, T. (2010), 'Neural activity associated with monitoring the oscillating threat value of a tarantula', *Proceedings of the National Academy of Sciences* **107**(47), 20582–20586.

Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. and Onnela, J.-P. (2010), 'Community structure in time-dependent, multiscale, and multiplex networks', *science* **328**(5980), 876–878.

Mueller, S., Wang, D., Fox, M. D., Yeo, B. T., Sepulcre, J., Sabuncu, M. R., Shafee, R., Lu, J. and Liu, H. (2013), 'Individual variability in functional connectivity architecture of the human brain', *Neuron* **77**(3), 586–595.

Nacewicz, B., Alexander, A., Kalin, N. and Davidson, R. (2014), The neurochemical underpinnings of human amygdala volume including subregional contributions., *in* 'Annual meeting Society of Biological Psychiatry', Society of Biological Psychiatry.

Najafi, M., Kinnison, J. and Pessoa, L. (2017), 'Intersubject brain network organization during dynamic anxious anticipation', *bioRxiv* p. 120451.

Najafi, M., McMenamin, B. and Pessoa, L. (2015), Large-scale brain networks are better characterized via overlapping communities, *in* 'Pattern Recognition in NeuroImaging (PRNI), 2015 International Workshop on', IEEE.

Najafi, M., McMenamin, B. W., Simon, J. Z. and Pessoa, L. (2016), 'Overlapping communities reveal rich structure in large-scale brain networks during rest and task conditions', *NeuroImage* **135**, 92–106.

Najafi, M. and Pessoa, L. (2016), Studying intersubject networks and standard graph measures during dynamic threat processing, *in* 'Program No. 36904/NNN442016 Neuroscience Meeting Planner', Vol. Online, Society for Neuroscience.

Nepusz, T., Petróczi, A., Négyessy, L. and Bazsó, F. (2008), 'Fuzzy communities and the concept of bridgeness in complex networks', *Physical Review E* **77**(1), 016107.

Neter, J., Kutner, M. H., Nachtsheim, C. J. and Wasserman, W. (1996), *Applied linear statistical models*, Vol. 4, Irwin Chicago.

Newman, M. E. (2006), 'Modularity and community structure in networks', *Proceedings of the national academy of sciences* **103**(23), 8577–8582.

Newman, M. E. (2010), *Networks: an introduction*, Oxford University Press, Oxford.

Nummenmaa, L., Glerean, E., Viinikainen, M., Jääskeläinen, I. P., Hari, R. and Sams, M. (2012), 'Emotions promote social interaction by synchronizing brain activity across individuals', *Proceedings of the National Academy of Sciences* **109**(24), 9599–9604.

Nummenmaa, L., Saarimäki, H., Glerean, E., Gotsopoulos, A., Jääskeläinen, I. P., Hari, R. and Sams, M. (2014), 'Emotional speech synchronizes brains across listeners and engages large-scale dynamic brain networks', *NeuroImage* **102**, 498–509.

Padmala, S. and Pessoa, L. (2011), 'Reward reduces conflict by enhancing attentional control and biasing visual cortical processing', *Journal of cognitive neuroscience* **23**(11), 3419–3432.

Palla, G., Derényi, I., Farkas, I. and Vicsek, T. (2005), 'Uncovering the overlapping community structure of complex networks in nature and society', *Nature* **435**(7043), 814–818.

Passingham, R. E., Stephan, K. E. and Kötter, R. (2002), 'The anatomical basis of functional localization in the cortex', *Nature Reviews Neuroscience* **3**(8), 606–616.

Pessoa, L. (2013), *The cognitive-emotional brain: From interactions to integration*, MIT press.

Pessoa, L. (2014), 'Understanding brain networks and brain organization', *Physics of life reviews* **11**(3), 400–435.

Pessoa, L. (2015), 'Précis on the cognitive-emotional brain', *Behavioral and Brain Sciences* **38**, e71.

Pessoa, L. (2017), 'A network model of the emotional brain', *Trends in Cognitive Sciences* .

Pessoa, L., McKenna, M., Gutierrez, E. and Ungerleider, L. (2002), 'Neural processing of emotional faces requires attention', *Proceedings of the National Academy of Sciences* **99**(17), 11458–11463.

Pessoa, L. and McMenamin, B. (2016), 'Dynamic networks in the emotional brain', *The Neuroscientist* p. 1073858416671936.

Pessoa, L. and Najafi, M. (2015), 'Complex-system causality in large-scale brain networks: Comment on foundational perspectives on causality in large-scale brain networks by m. mannino and sl bressler', *Physics of life reviews* **15**, 124–127.

Poldrack, R. A. (2006), 'Can cognitive processes be inferred from neuroimaging data?', *Trends in cognitive sciences* **10**(2), 59–63.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., Nichols, T. E., Poline, J.-B., Vul, E. and Yarkoni, T. (2017), 'Scanning the horizon: towards transparent and reproducible neuroimaging research', *Nature Reviews Neuroscience* .

Poldrack, R. A., Halchenko, Y. O. and Hanson, S. J. (2009), 'Decoding the large-scale structure of brain function by classifying mental states across individuals', *Psychological Science* **20**(11), 1364–1372.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. and Petersen, S. E. (2012), 'Spurious but systematic correlations in functional connectivity mri networks arise from subject motion', *Neuroimage* **59**(3), 2142–2154.

Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N. and Petersen, S. E. (2013), 'Evidence for hubs in human functional brain networks', *Neuron* **79**(4), 798–813.

Pruessner, J. C., Dedovic, K., Khalili-Mahani, N., Engert, V., Pruessner, M., Buss, C., Renwick, R., Dagher, A., Meaney, M. J. and Lupien, S. (2008), 'Deactivation of the limbic system during acute psychosocial stress: evidence from positron emission tomography and functional magnetic resonance imaging studies', *Biological psychiatry* **63**(2), 234–240.

Purves, D., Cabeza, R., Huettel, S. A., LaBar, K. S., Platt, M. L., Woldorff, M. G. and Brannon, E. M. (2008), *Cognitive Neuroscience*, Sunderland: Sinauer Associates, Inc.

Rakic, P., Bourgeois, J.-P., Eckenhoff, M. F., Zecevic, N. and Goldman-Rakic, P. S. (1986), 'Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex', *Science* **232**, 232–236.

Rashid, B., Damaraju, E., Pearlson, G. D. and Calhoun, V. D. (2014), 'Dynamic connectivity states estimated from resting fmri identify differences among schizophrenia, bipolar disorder, and healthy control subjects', *Frontiers in human neuroscience* **8**, 897.

Riedel, M. C., Ray, K. L., Dick, A. S., Sutherland, M. T., Hernandez, Z., Fox, P. M., Eickhoff, S. B., Fox, P. T. and Laird, A. R. (2015), 'Meta-analytic connectivity and behavioral parcellation of the human cerebellum', *Neuroimage* **117**, 327–342.

Rosvall, M. and Bergstrom, C. T. (2008), 'Maps of random walks on complex networks reveal community structure', *Proceedings of the National Academy of Sciences* **105**(4), 1118–1123.

Roy, M., Shohamy, D., Daw, N., Jepma, M., Wimmer, G. E. and Wager, T. D. (2014), 'Representation of aversive prediction errors in the human periaqueductal gray', *Nature neuroscience* **17**(11), 1607–1612.

Rubinov, M. and Sporns, O. (2011), 'Weight-conserving characterization of complex functional brain networks', *Neuroimage* **56**(4), 2068–2079.

Saad, Z. S. and Reynolds, R. C. (2012), 'Suma', *Neuroimage* **62**(2), 768–773.

Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L. and Smith, S. M. (2014), 'Automatic denoising of functional mri data: combining independent component analysis and hierarchical fusion of classifiers', *Neuroimage* **90**, 449–468.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L. and Greicius, M. D. (2007), 'Dissociable intrinsic connectivity networks for salience processing and executive control', *Journal of Neuroscience* **27**(9), 2349–2356.

Seghier, M. L. (2013), 'The angular gyrus multiple functions and multiple subdivisions', *The Neuroscientist* **19**(1), 43–61.

Shackman, A. J. and Fox, A. S. (2016), 'Contributions of the central extended amygdala to fear and anxietycontributions of the central extended amygdala to fear and anxiety', *Journal of neuroscience* **36**(31), 8050–8063.

Shallice, T. (1988), *From neuropsychology to mental structure*, Cambridge University Press.

Shallice, T. and Cooper, R. P. (2012), 'The organisation of mind', *cortex* **48**(10), 1366–1370.

Shannon, C. E. (2001), 'A mathematical theory of communication', *ACM SIGMOBILE Mobile Computing and Communications Review* **5**(1), 3–55.

Shattuck, D. W. and Leahy, R. M. (2002), 'Brainsuite: an automated cortical surface identification tool', *Medical image analysis* **6**(2), 129–142.

Sherman, S. M. and Guillery, R. W. (2013), *Functional connections of cortical areas: a new view from the thalamus*, MIT Press.

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A. and Hasson, U. (2016), 'Dynamic reconfiguration of the default mode network during narrative comprehension', *Nature communications* **7**.

Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D. A., Griffanti, L., Harms, M. P. et al. (2013), 'Resting-state fmri in the human connectome project', *Neuroimage* **80**, 144–168.

Smith, S. M., Miller, K. L., Moeller, S., Xu, J., Auerbach, E. J., Woolrich, M. W., Beckmann, C. F., Jenkinson, M., Andersson, J., Glasser, M. F. et al. (2012), 'Temporally-independent functional modes of spontaneous brain activity', *Proceedings of the National Academy of Sciences* **109**(8), 3131–3136.

Somerville, L. H., Whalen, P. J. and Kelley, W. M. (2010), 'Human bed nucleus of the stria terminalis indexes hypervigilant threat monitoring', *Biological psychiatry* **68**(5), 416–424.

Sporns, O. (2013), 'Structure and function of complex brain networks', *Dialogues Clin Neurosci* **15**(3), 247–262.

Sreenivasan, K. K., Curtis, C. E. and DEsposito, M. (2014), 'Revisiting the role of persistent neural activity during working memory', *Trends in cognitive sciences* **18**(2), 82–89.

Stephens, G. J., Silbert, L. J. and Hasson, U. (2010), 'Speaker–listener neural coupling underlies successful communication', *Proceedings of the National Academy of Sciences* **107**(32), 14425–14430.

Theiss, J. D., Ridgewell, C., McHugo, M., Heckers, S. and Blackford, J. U. (2017), 'Manual segmentation of the human bed nucleus of the stria terminalis using 3t mri', *NeuroImage* **146**, 288–292.

Tomasi, D. and Volkow, N. D. (2011), 'Functional connectivity hubs in the human brain', *Neuroimage* **57**(3), 908–917.

Torrisi, S., O'connell, K., Davis, A., Reynolds, R., Balderston, N., Fudge, J. L., Grillon, C. and Ernst, M. (2015), 'Resting state connectivity of the bed nucleus of the stria terminalis at ultra-high field', *Human brain mapping* **36**(10), 4076–4088.

Tovote, P., Fadok, J. P. and Lüthi, A. (2015), 'Neuronal circuits for fear and anxiety', *Nature Reviews Neuroscience* **16**(6), 317–331.

Van Dijk, K. R., Sabuncu, M. R. and Buckner, R. L. (2012), 'The influence of head motion on intrinsic functional connectivity mri', *Neuroimage* **59**(1), 431–438.

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., Consortium, W.-M. H. et al. (2013), 'The wu-minn human connectome project: an overview', *Neuroimage* **80**, 62–79.

Wager, T. D., Barrett, L. F., Bliss-Moreau, E., Lindquist, K., Duncan, S., Kober, H., Joseph, J., Davidson, M. and Mize, J. (2008), 'The neuroimaging of emotion', *Handbook of emotions* **3**, 249–271.

Wager, T. D., Waugh, C. E., Lindquist, M., Noll, D. C., Fredrickson, B. L. and Taylor, S. F. (2009), 'Brain mediators of cardiovascular responses to social threat: part i: Reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity', *Neuroimage* **47**(3), 821–835.

Wasserman, S. and Faust, K. (1994), *Social network analysis: Methods and applications*, Vol. 8, Cambridge university press.

Wu, Y., Hernández-Lobato, J. M. and Ghahramani, Z. (2013), Dynamic covariance models for multivariate financial time series., *in* 'ICML (3)', pp. 558–566.

Xie, J., Kelley, S. and Szymanski, B. K. (2013), 'Overlapping community detection in networks: The state-of-the-art and comparative study', *Acm computing surveys (csur)* **45**(4), 43.

Yeo, B. T., Krienen, F. M., Chee, M. W. and Buckner, R. L. (2014), 'Estimates of segregation and overlap of functional connectivity networks in the human cerebral cortex', *Neuroimage* **88**, 212–227.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R. et al. (2011), 'The organization of the human cerebral cortex estimated by intrinsic functional connectivity', *Journal of neurophysiology* **106**(3), 1125–1165.

Yu, H., Kim, P. M., Sprecher, E., Trifonov, V. and Gerstein, M. (2007), 'The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics', *PLoS Comput Biol* **3**(4), e59.